



UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

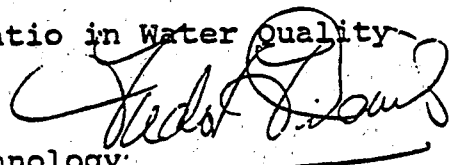
FEB 22 1994

OFFICE OF
WATER

EPA-823-B-94-001

MEMORANDUM

SUBJECT: Use of the Water-Effect Ratio in Water Quality Standards

FROM: Tudor T. Davies, Director
Office of Science and Technology 

TO: Water Management Division Directors, Regions I - X
State Water Quality Standards Program Directors

PURPOSE

There are two purposes for this memorandum.

The first is to transmit the Interim Guidance on the Determination and Use of Water-Effect Ratios for Metals. EPA committed to developing this guidance to support implementation of federal standards for those States included in the National Toxics Rule.

The second is to provide policy guidance on whether a State's application of a water-effect ratio is a site-specific criterion adjustment subject to EPA review and approval/disapproval.

BACKGROUND

In the early 1980's, members of the regulated community expressed concern that EPA's laboratory-derived water quality criteria might not accurately reflect site-specific conditions because of the effects of water chemistry and the ability of species to adapt over time. In response to these concerns, EPA created three procedures to derive site-specific criteria. These procedures were published in the Water Quality Standards Handbook, 1983.



Site-specific criteria are allowed by regulation and are subject to EPA review and approval. The Federal water quality standards regulation at section 131.11(b)(1) provides States with the opportunity to adopt water quality criteria that are "...modified to reflect site-specific conditions." Under section 131.5(a)(2), EPA reviews standards to determine "whether a State has adopted criteria to protect the designated water uses."

On December 22, 1992, EPA promulgated the National Toxics Rule which established Federal water quality standards for 14 States which had not met the requirements of Clean Water Act Section 303(c)(2)(B). As part of that rule, EPA gave the States discretion to adjust the aquatic life criteria for metals to reflect site-specific conditions through use of a water-effect ratio. A water-effect ratio is a means to account for a difference between the toxicity of the metal in laboratory dilution water and its toxicity in the water at the site.

In promulgating the National Toxics Rule, EPA committed to issuing updated guidance on the derivation of water-effect ratios. The guidance reflects new information since the previous guidance and is more comprehensive in order to provide greater clarity and increased understanding. This new guidance should help standardize procedures for deriving water-effect ratios and make results more comparable and defensible.

Recently, an issue arose concerning the most appropriate form of metals upon which to base water quality standards. On October 1, 1993, EPA issued guidance on this issue which indicated that measuring the dissolved form of metal is the recommended approach. This new policy however, is prospective and does not affect the criteria in the National Toxics Rule. Dissolved metals criteria are not generally numerically equal to total recoverable criteria and the October 1, 1993 guidance contains recommendations for correction factors for fresh water criteria. The determination of site-specific criteria is applicable to criteria expressed as either total recoverable metal or as dissolved metal.

DISCUSSION

Existing guidance and practice are that EPA will approve site-specific criteria developed using appropriate procedures. That policy continues for the options set forth in the interim guidance transmitted today, regardless of whether the resulting criterion is equal to or more or less stringent than the EPA national 304(a) guidance. This interim guidance supersedes all guidance concerning water-effect ratios previously issued by the Agency.

Each of the three options for deriving a final water-effect ratio presented in this interim guidance meets the scientific and technical acceptability test for deriving site-specific criteria.

Option 3 is the simplest, least restrictive and generally the least expensive approach for situations where simulated downstream water appropriately represents a "site." It is a fully acceptable approach for deriving the water-effect ratio although it will generally provide a lower water-effect ratio than the other 2 options. The other 2 options may be more costly and time consuming if more than 3 sample periods and water-effect ratio measurements are made, but are more accurate, and may yield a larger, but more scientifically defensible site specific criterion.

Site-specific criteria, properly determined, will fully protect existing uses. The waterbody or segment thereof to which the site-specific criteria apply must be clearly defined. A site can be defined by the State and can be any size, small or large, including a watershed or basin. However, the site-specific criteria must protect the site as a whole. It is likely to be more cost-effective to derive any site-specific criteria for as large an area as possible or appropriate. It is emphasized that site-specific criteria are ambient water quality criteria applicable to a site. They are not intended to be direct modifications to National Pollutant Discharge Elimination System (NPDES) permit limits. In most cases the "site" will be synonymous with a State's "segment" in its water quality standards. By defining sites on a larger scale, multiple dischargers can collaborate on water-effect ratio testing and attain appropriate site-specific criteria at a reduced cost.

More attention has been given to water-effect ratios recently because of the numerous discussions and meetings on the entire question of metals policy and because WERs were specifically applied in the National Toxics Rule. In comments on the proposed National Toxics Rule, the public questioned whether the EPA promulgation should be based solely on the total recoverable form of a metal. For the reasons set forth in the final preamble, EPA chose to promulgate the criteria based on the total recoverable form with a provision for the application of a water-effect ratio. In addition, this approach was chosen because of the unique difficulties of attempting to authorize site-specific criteria modifications for nationally promulgated criteria.

EPA now recommends the use of dissolved metals for States revising their water quality standards. Dissolved criteria may also be modified by a site-specific adjustment.

While the regulatory application of the water-effect ratio applied only to the 10 jurisdictions included in the final National Toxics Rule for aquatic life metals criteria, we understood that other States would be interested in applying WERs to their adopted water quality standards. The guidance upon which to base the judgment of the acceptability of the water-effect ratio applied by the State is contained in the attached Interim Guidance on The Determination and Use of Water-Effect Ratios for Metals. It should be noted that this guidance also provides additional information on the recalculation procedure for site-specific criteria modifications.

Status of the Water-effect Ratio (WER) in non-National Toxics Rule States

A central question concerning WERs is whether their use by a State results in a site-specific criterion subject to EPA review and approval under Section 303(c) of the Clean Water Act?

Derivation of a water-effect ratio by a State is a site-specific criterion adjustment subject to EPA review and approval/disapproval under Section 303(c). There are two options by which this review can be accomplished.

Option 1: A State may derive and submit each individual water-effect ratio determination to EPA for review and approval. This would be accomplished through the normal review and revision process used by a State.

Option 2: A State can amend its water quality standards to provide a formal procedure which includes derivation of water-effect ratios, appropriate definition of sites, and enforceable monitoring provisions to assure that designated uses are protected. Both this procedure and the resulting criteria would be subject to full public participation requirements. Public review of a site-specific criterion could be accomplished in conjunction with the public review required for permit issuance. EPA would review and approve/disapprove this protocol as a revised standard once. For public information, we recommend that once a year the State publish a list of site-specific criteria.

An exception to this policy applies to the waters of the jurisdictions included in the National Toxics Rule. The EPA review is not required for the jurisdictions included in the National Toxics Rule where EPA established the procedure for the State for application to the criteria promulgated. The National Toxics Rule was a formal rulemaking process with notice and comment by which EPA pre-authorized the use of a correctly applied water-effect ratio. That same process has not yet taken place in States not included in the National Toxics Rule.

However, the National Toxics Rule does not affect State authority to establish scientifically defensible procedures to determine Federally authorized WERs, to certify those WERs in NPDES permit proceedings, or to deny their application based on the State's risk management analysis.

As described in Section 131.36(b)(iii) of the water quality standards regulation (the official regulatory reference to the National Toxics Rule), the water-effect ratio is a site-specific calculation. As indicated on page 60866 of the preamble to the National Toxics Rule, the rule was constructed as a rebuttable presumption. The water-effect ratio is assigned a value of 1.0 until a different water-effect ratio is derived from suitable tests representative of conditions in the affected waterbody. It is the responsibility of the State to determine whether to rebut the assumed value of 1.0 in the National Toxics Rule and apply another value of the water-effect ratio in order to establish a site-specific criterion. The site-specific criterion is then used to develop appropriate NPDES permit limits. The rule thus provides a State with the flexibility to derive an appropriate site-specific criterion for specific waterbodies.

As a point of emphasis, although a water-effect ratio affects permit limits for individual dischargers, it is the State in all cases that determines if derivation of a site-specific criterion based on the water-effect ratio is allowed and it is the State that ensures that the calculations and data analysis are done completely and correctly.

CONCLUSION

This interim guidance explains and clarifies the use of site-specific criteria. It is issued as interim guidance because it will be included as part of the process underway for review and possible revision of the national aquatic life criteria development methodology guidelines. As part of that review, this interim guidance is subject to amendment based on comments, especially those from the users of the guidance. At the end of the guidelines revision process the guidance will be issued as "final."

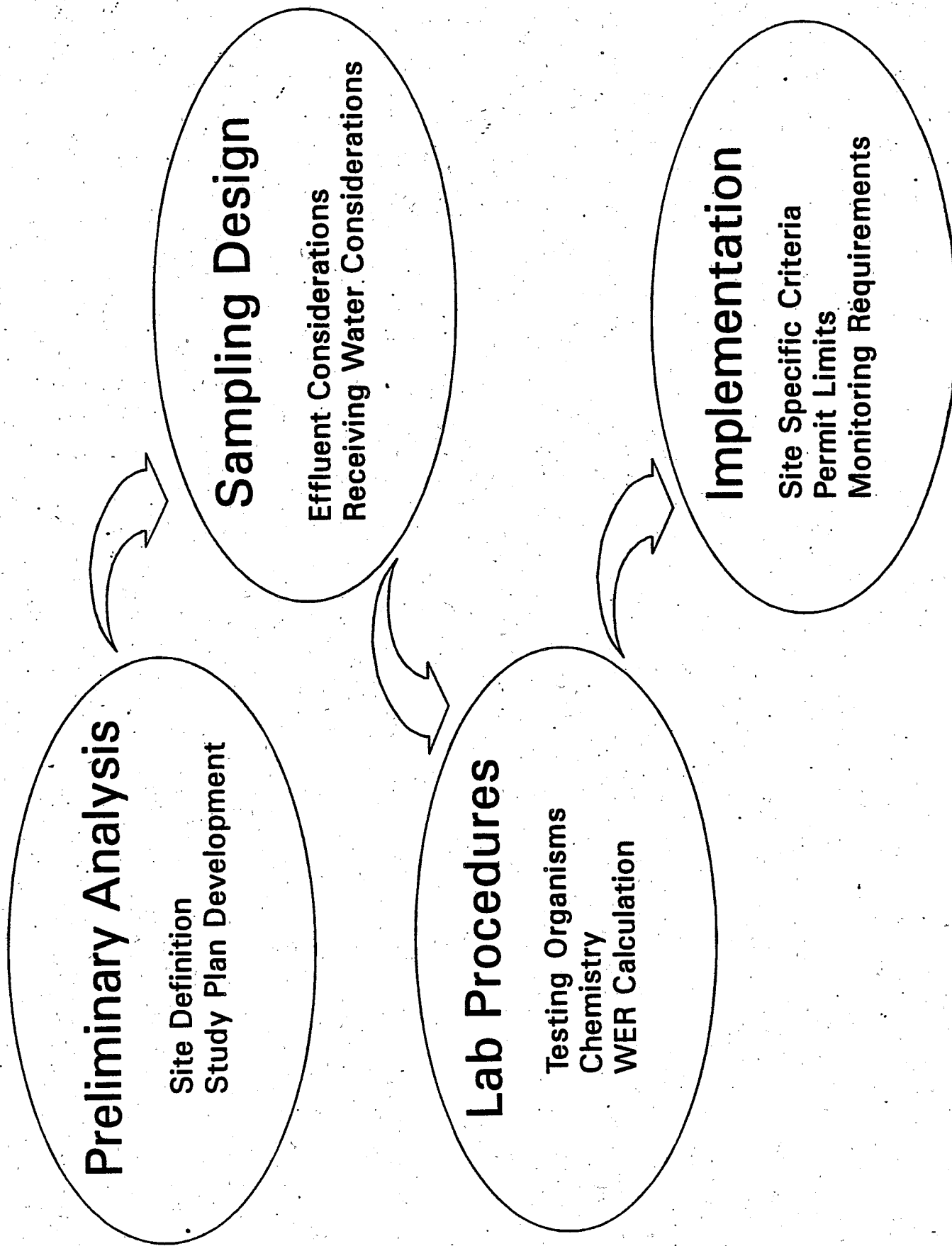
EPA is interested in and encourages the submittal of high quality datasets that can be used to provide insights into the use of these guidelines and procedures. Such data and technical comments should be submitted to Charles E. Stephan at EPA's Environmental Research Laboratory at Duluth, MN. A complete address, telephone number and fax number for Mr. Stephan are included in the guidance itself. Other questions or comments should be directed to the Standards and Applied Science Division (mail code 4305, telephone 202-260-1315).

There is attached to this memorandum a simplified flow diagram and an implementation procedure. These are intended to aid a user by placing the water-effect ratio procedure in the context of proceeding from a site-specific criterion to a permit limit. Following these attachments is the guidance itself.

Attachments

cc: Robert Perciasepe, OW
Martha G. Prothro, OW
William Diamond, SASD
Margaret Stasikowski, HECD
Mike Cook, OWEC
Cynthia Dougherty, OWEC
Lee Schroer, OGC
Susan Lepow, OGC
Courtney Riordan, ORD
ORD (Duluth and Narragansett Laboratories)
ESD Directors, Regions I - VIII, X
ESD Branch, Region IX
Water Quality Standards Coordinators, Regions I - X

WER Implementation



WATER-EFFECT RATIO IMPLEMENTATION

PRELIMINARY ANALYSIS & PLAN FORMULATION

- Site definition

- How many discharges must be accounted for? Tributaries? See page 17.
- What is the waterbody type? (i.e., stream, tidal river, bay, etc.). See page 44 and Appendix A.
- How can these considerations best be combined to define the relevant geographic "site"? See Appendix A @ page 82.

- Plan Development for Regulatory Agency Review

- Is WER method 1 or 2 appropriate? (e.g., Is design flow a meaningful concept or are other considerations paramount?). See page 6.
- Define the effluent & receiving water sample locations
- Describe the temporal sample collection protocols proposed. See page 48.
- Can simulated site water procedure be done, or is downstream sampling required? See Appendix A.
- Describe the testing protocols - test species, test type, test length, etc. See page 45, 50; Appendix I.
- Describe the chemical testing proposed. See Appendix C.
- Describe other details of study - flow measurement, QA/QC, number of sampling periods proposed, to whom the results are expected to apply, schedule, etc.

SAMPLING DESIGN FOR STREAMS

- Discuss the quantification of the design streamflow (e.g., 7Q10) - USGS gage directly, by extrapolation from USGS gage, or ?

- Effluents

- measure flows to determine average for sampling day
- collect 24 hour composite using "clean" equipment and appropriate procedures; avoid the use of the plant's daily composite sample as a shortcut.

- Streams

- measure flow (use current meter or read from gage if available) to determine dilution with effluent; and to check if within acceptable range for use of the data (i.e., design flow to 10 times the design flow).
- collect 24 hour composite of upstream water.

LABORATORY PROCEDURES (NOTE: These are described in detail in interim guidance).

- Select appropriate primary & secondary tests
- Determine appropriate cmcWER and/or cccWER
- Perform chemistry using clean procedures, with methods that have adequate sensitivity to measure low concentrations, and use appropriate QA/QC
- Calculate final water-effect ratio (FWER) for site. See page 36.

IMPLEMENTATION

- Assign FWERS and the site specific criteria for each metal to each discharger (if more than one).
- perform a waste load allocation and total maximum daily load (if appropriate) so that each discharger is provided a permit limit.
- establish monitoring condition for periodic evaluation of instream biology (recommended)
- establish a permit condition for periodic testing of WER to verify site-specific criterion (NTR recommendation)

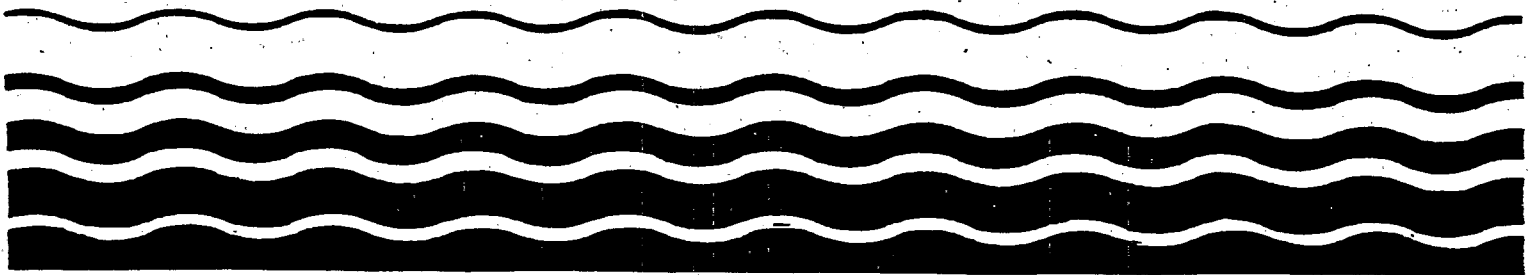
United States
Environmental Protection
Agency

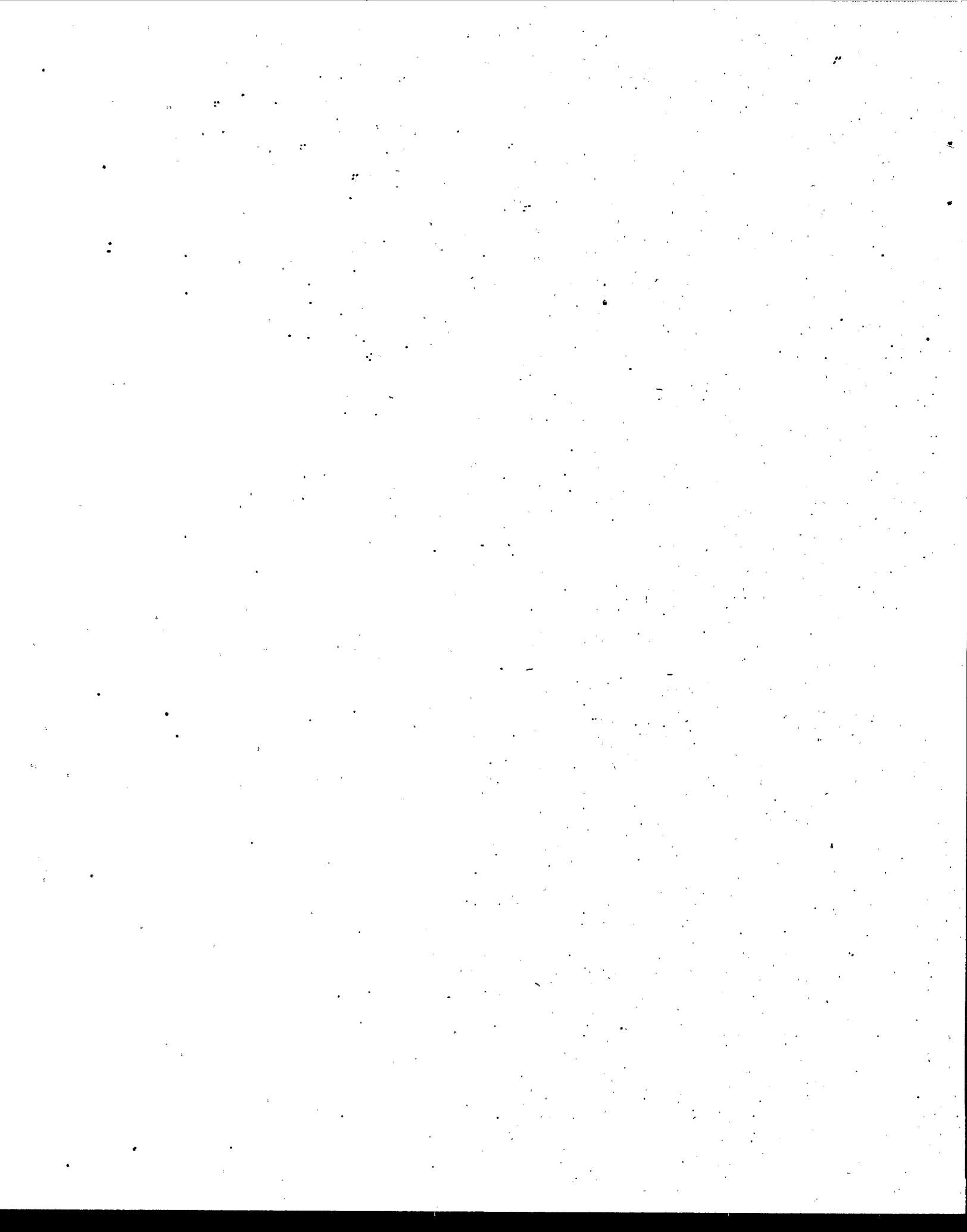
Office of Water
Office of Science & Technology
(Mail Code 4305)

February 1994
EPA-823-B-94-001



Interim Guidance on Determination and Use of Water-Effect Ratios for Metals





**Interim Guidance on
Determination and Use of
Water-Effect Ratios for Metals**

February 1994

U.S. Environmental Protection Agency

**Office of Water
Office of Science and Technology
Washington, D.C.**

**Office of Research and Development
Environmental Research Laboratories
Duluth, Minnesota
Narragansett, Rhode Island**

NOTICES

This document has been reviewed by the Environmental Research Laboratories, Duluth, MN and Narragansett, RI (Office of Research and Development) and the Office of Science and Technology (Office of Water), U.S. Environmental Protection Agency, and approved for publication.

Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

FOREWORD

This document provides interim guidance concerning the experimental determination of water-effect ratios (WERs) for metals; some aspects of the use of WERs are also addressed. It is issued in support of EPA regulations and policy initiatives involving the application of water quality criteria and standards for metals. This document is agency guidance only. It does not establish or affect legal rights or obligations. It does not establish a binding norm or prohibit alternatives not included in the document. It is not finally determinative of the issues addressed. Agency decisions in any particular case will be made by applying the law and regulations on the basis of specific facts when regulations are promulgated or permits are issued.

This document is expected to be revised periodically to reflect advances in this rapidly evolving area. Comments, especially those accompanied by supporting data, are welcomed and should be sent to: Charles E. Stephan, U.S. EPA, 6201 Congdon Boulevard, Duluth MN 55804 (TEL: 218-720-5510; FAX: 218-720-5539).



UNITED STATES ENVIRONMENTAL PROTECTION AGENCY
WASHINGTON, D.C. 20460

FEB 22 1994

OFFICE OF
WATER

OFFICE OF SCIENCE AND TECHNOLOGY POSITION STATEMENT

Section 131.11(b)(ii) of the water quality standards regulation (40 CFR Part 131) provides the regulatory mechanism for a State to develop site-specific criteria for use in water quality standards. Adopting site-specific criteria in water quality standards is a State option--not a requirement. The Environmental Protection Agency (EPA) in 1983 provided guidance on scientifically acceptable methods by which site-specific criteria could be developed.

The interim guidance provided in this document supersedes all guidance concerning water-effect ratios and the Indicator Species Procedure given in Chapter 4 of the Water Quality Standards Handbook issued by EPA in 1983 and in Guidelines for Deriving Numerical Aquatic Site-Specific Water Quality Criteria by Modifying National Criteria, 1984. Appendix B also supersedes the guidance in these earlier documents for the Recalculation Procedure for performing site-specific criteria modifications.

This interim guidance fulfills a commitment made in the final rule to establish numeric criteria for priority toxic pollutants (57 FR 60848, December 22, 1992, also known as the "National Toxics Rule"). This guidance also is applicable to pollutants other than metals with appropriate modifications, principally to chemical analyses.

Except for the jurisdictions subject to the aquatic life criteria in the national toxics rule, water-effect ratios are site-specific criteria subject to review and approval by the appropriate EPA Regional Administrator. Site-specific criteria are new or revised criteria subject to the normal EPA review requirements established in Clean Water Act § 303(c). For the States in the National Toxics Rule, EPA has established that site-specific water-effect ratios may be applied to the criteria promulgated in the rule to establish site-specific criteria. The water-effect ratio portion of these criteria would still be subject to State review before the development of total maximum daily loads, waste load allocations or translation into NPDES permit limits. EPA would only review these water-effect ratios during its oversight review of these State programs or review of State-issued permits.

Each of the three options for deriving a final water-effect ratio presented on page 36 of this interim guidance meets the scientific and technical acceptability test for deriving site-specific criteria specified in the water quality standards regulation (40 CFR 131.11(a)). Option 3 is the simplest, least restrictive and generally the least expensive approach for situations where simulated downstream water appropriately represents a "site." Option 3 requires experimental determination of three water-effect ratios with the primary test species that are determined during any season (as long as the downstream flow is between 2 and 10 times design flow conditions.) The final WER is generally (but not always) the lowest experimentally determined WER. Deriving a final water-effect ratio using option 3 with the use of simulated downstream water for a situation where this simulation appropriately represents a "site", is a fully acceptable approach for deriving a water-effect ratio for use in determining a site-specific criterion, although it will generally provide a lower water-effect ratio than the other 2 options.

As indicated in the introduction to this guidance, the determination of a water-effect ratio may require substantial resources. A discharger should consider cost-effective, preliminary measures described in this guidance (e.g., use of "clean" sampling and chemical analytical techniques or in non-NTR States, a recalculated criterion) to determine if an indicator species site-specific criterion is really needed. It may be that an appropriate site-specific criterion is actually being attained. In many instances, use of these other measures may eliminate the need for deriving final water-effect ratios. The methods described in this interim guidance should be sufficient to develop site-specific criteria that resolve concerns of dischargers when there appears to be no instream toxicity from a metal but, where (a) a discharge appears to exceed existing or proposed water quality-based permit limits, or (b) an instream concentration appears to exceed an existing or proposed water quality criterion.

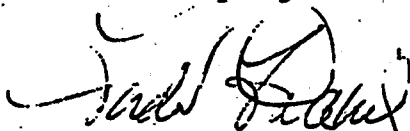
This guidance describes 2 different methods for determining water-effect ratios. Method 1 has 3 options each of which may only require 3 sampling periods. However options 1 and 2 may be expanded and require a much greater effort. While this position statement has discussed the simplest, least expensive option for method 1 (the single discharge to a stream) to illustrate that site specific criteria are feasible even when only small dischargers are affected, water-effect ratios may be calculated using any of the other options described in the guidance if the State/discharger believe that there is reason to expect that a more accurate site-specific criterion will result from the increased cost and complexity inherent in conducting the

additional tests and analyzing the results. Situations where this could be the case include, for example, where seasonal effects in receiving water quality or in discharge quality need to be assessed.

In addition, EPA will consider other scientifically defensible approaches in developing final water-effect ratios as authorized in 40 CFR 131.11. However, EPA strongly recommends that before a State/discharger implements any approach other than one described in this interim guidance, discussions be held with appropriate EPA regional offices and Office of Research and Development's scientists before actual testing begins. These discussions would be to ensure that time and resources are not wasted on scientifically and technically unacceptable approaches. It remains EPA's responsibility to make final decisions on the scientific and technical validity of alternative approaches to developing site-specific water quality criteria.

EPA is fully cognizant of the continuing debate between what constitutes guidance and what is a regulatory requirement. Developing site-specific criteria is a State regulatory option. Using the methodology correctly as described in this guidance assures the State that EPA will accept the result. Other approaches are possible and logically should be discussed with EPA prior to implementation.

The Office of Science and Technology believes that this interim guidance advances the science of determining site-specific criteria and provides policy guidance that States and EPA can use in this complex area. It reflects the scientific advances in the past 10 years and the experience gained from dealing with these issues in real world situations. This guidance will help improve implementation of water quality standards and be the basis for future progress.



Tudor T. Davies, Director
Office of Science And Technology
Office of Water

CONTENTS

	<u>Page</u>
Notices	ii
Foreword	iii
Office of Science and Technology Position Statement	iv
Appendices	viii
Figures	ix
Acknowledgments	x
Executive Summary	xi
Abbreviations	xiii
Glossary	xiv
Preface	xvi
 Introduction	 1
 Method 1	 17
A. Experimental Design	17
B. Background Information and Initial Decisions	44
C. Selecting Primary and Secondary Tests	45
D. Acquiring and Acclimating Test Organisms	47
E. Collecting and Handling Upstream Water and Effluent	48
F. Laboratory Dilution Water	49
G. Conducting Tests	50
H. Chemical and Other Measurements	55
I. Calculating and Interpreting the Results	57
J. Reporting the Results	62
 Method 2	 65
 References	 76

APPENDICES

	<u>Page</u>
A. Comparison of WERs Determined Using Upstream and Downstream Water	79
B. The Recalculation Procedure	90
C. Guidance Concerning the Use of "Clean Techniques" and QA/QC when Measuring Trace Metals	98
D. Relationships between WERs and the Chemistry and Toxicology of Metals	109
E. U.S. EPA Aquatic Life Criteria Documents for Metals	134
F. Considerations Concerning Multiple-Metal, Multiple-Discharge, and Special Flowing-Water Situations	135
G. Additivity and the Two Components of a WER Determined Using Downstream Water	139
H. Special Considerations Concerning the Determination of WERs with Saltwater Species	145
I. Suggested Toxicity Tests for Determining WERs for Metals	147
J. Recommended Salts of Metals	153

FIGURES

	<u>Page</u>
1. Four Ways to Derive a Permit Limit	16
2. Calculating an Adjusted Geometric Mean	71
3. An Example Derivation of a FWER	72
4. Reducing the Impact of Experimental Variation	73
5. Calculating an LC50 (or EC50) by Interpolation	74
6. Calculating a Time-Weighted Average	75
B1. An Example of the Deletion Process Using Three Phyla . . .	97
D1. A Scheme for Classifying Forms of Metal in Water	111
D2. An Example of the Empirical Extrapolation Process	125
D3. The Internal Consistency of the Two Approaches	126
D4. The Application of the Two Approaches	128
D5. A Generalized Complexation Curve	131
D6. A Generalized Precipitation Curve	132

ACKNOWLEDGMENTS

This document was written by:

Charles E. Stephan, U.S. EPA, ORD, Environmental Research Laboratory, Duluth, MN.

William H. Peltier, U.S. EPA, Region IV, Environmental Services Division, Athens, GA.

David J. Hansen, U.S. EPA, ORD, Environmental Research Laboratory, Narragansett, RI.

Charles G. Delos, U.S. EPA, Office of Water, Health and Ecological Criteria Division, Washington, DC.

Gary A. Chapman, U.S. EPA, ORD, Environmental Research Laboratory (Narragansett), Pacific Ecosystems Branch, Newport, OR.

The authors thank all the people who participated in the open discussion of the experimental determination of water-effect ratios on Tuesday evening, January 26, 1993 in Annapolis, MD. Special thanks go to Herb Allen, Bill Beckwith, Ken Bruland, Lee Dunbar, Russ Erickson, and Carlton Hunt for their technical input on this project, although none of them necessarily agree with everything in this document. Comments by Kent Ballentine, Karen Gourdine, Mark Hicks, Suzanne Lussier, Nelson Thomas, Bob Spehar, Fritz Wagener, Robb Wood, and Phil Woods on various drafts, or portions of drafts, were also very helpful, as were discussions with several other individuals.

EXECUTIVE SUMMARY

A variety of physical and chemical characteristics of both the water and the metal can influence the toxicity of a metal to aquatic organisms in a surface water. When a site-specific aquatic life criterion is derived for a metal, an adjustment procedure based on the toxicological determination of a water-effect ratio (WER) may be used to account for a difference between the toxicity of the metal in laboratory dilution water and its toxicity in the water at the site. If there is a difference in toxicity and it is not taken into account, the aquatic life criterion for the body of water will be more or less protective than intended by EPA's Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses. After a WER is determined for a site, a site-specific aquatic life criterion can be calculated by multiplying an appropriate national, state, or recalculated criterion by the WER. Most WERs are expected to be equal to or greater than 1.0, but some might be less than 1.0. Because most aquatic life criteria consist of two numbers, i.e., a Criterion Maximum Concentration (CMC) and a Criterion Continuous Concentration (CCC), either a cmcWER or a cccWER or both might be needed for a site. The cmcWER and the cccWER cannot be assumed to be equal, but it is not always necessary to determine both.

In order to determine a WER, side-by-side toxicity tests are performed to measure the toxicity of the metal in two dilution waters. One of the waters has to be a water that would be acceptable for use in laboratory toxicity tests conducted for the derivation of national water quality criteria for aquatic life. In most situations, the second dilution water will be a simulated downstream water that is prepared by mixing upstream water and effluent in an appropriate ratio; in other situations, the second dilution water will be a sample of the actual site water to which the site-specific criterion is to apply. The WER is calculated by dividing the endpoint obtained in the site water by the endpoint obtained in the laboratory dilution water. A WER should be determined using a toxicity test whose endpoint is close to, but not lower than, the CMC and/or CCC that is to be adjusted.

A total recoverable WER can be determined if the metal in both of the side-by-side toxicity tests is analyzed using the total recoverable measurement, and a dissolved WER can be determined if the metal is analyzed in both tests using the dissolved measurement. Thus four WERs can be determined:

- Total recoverable cmcWER.

- Total recoverable cccWER.

- Dissolved cmcWER.

- Dissolved cccWER.

A total recoverable WER is used to calculate a total recoverable site-specific criterion from a total recoverable national, state,

or recalculated aquatic life criterion, whereas a dissolved WER is used to calculate a dissolved site-specific criterion from a dissolved criterion. WERs are determined individually for each metal at each site; WERs cannot be extrapolated from one metal to another, one effluent to another, or one site water to another.

Because determining a WER requires substantial resources, the desirability of obtaining a WER should be carefully evaluated:

1. Determine whether use of "clean techniques" for collecting, handling, storing, preparing, and analyzing samples will eliminate the reason for considering determination of a WER, because existing data concerning concentrations of metals in effluents and surface waters might be erroneously high.
2. Evaluate the potential for reducing the discharge of the metal.
3. Investigate possible constraints on the permit limits, such as antibacksliding and antidegradation requirements and human health and wildlife criteria.
4. Consider use of the Recalculation Procedure.
5. Evaluate the cost-effectiveness of determining a WER.

If the determination of a WER is desirable, a detailed workplan for should be submitted to the appropriate regulatory authority (and possibly to the Water Management Division of the EPA Regional Office) for comment. After the workplan is completed, the initial phase should be implemented, the data should be evaluated, and the workplan should be revised if appropriate.

Two methods are used to determine WERs. Method 1, which is used to determine cccWERS that apply near plumes and to determine all cmcWERS, uses data concerning three or more distinctly separate sampling events. It is best if the sampling events occur during both low-flow and higher-flow periods. When sampling does not occur during both low and higher flows, the site-specific criterion is derived in a more conservative manner due to greater uncertainty. For each sampling event, a WER is determined using a selected toxicity test; for at least one of the sampling events, a confirmatory WER is determined using a different test.

Method 2, which is used to determine a cccWER for a large body of water outside the vicinities of plumes, requires substantial site-specific planning and more resources than Method 1. WERs are determined using samples of actual site water obtained at various times, locations, and depths to identify the range of WERs in the body of water. The WERs are used to determine how many site-specific CCCs should be derived for the body of water and what the one or more CCCs should be.

The guidance contained herein replaces previous agency guidance concerning (a) the determination of WERs for use in the derivation of site-specific aquatic life criteria for metals and (b) the Recalculation Procedure. This guidance is designed to apply to metals, but the principles apply to most pollutants.

ABBREVIATIONS

ACR:	Acute-Chronic Ratio
CCC:	Criterion Continuous Concentration
CMC:	Criterion Maximum Concentration
CRM:	Certified Reference Material
FAV:	Final Acute Value
FCV:	Final Chronic Value
FW:	Freshwater
FWER:	Final Water-Effect Ratio
GMAV:	Genus Mean Acute Value
HCME:	Highest Concentration of the Metal in the Effluent
MDR:	Minimum Data Requirement
NTR:	National Toxics Rule
QA/QC:	Quality Assurance/Quality Control
SMAV:	Species Mean Acute Value
SW:	Saltwater
TDS:	Total Dissolved Solids
TIE:	Toxicity Identification Evaluation
TMDL:	Total Maximum Daily Load
TOC:	Total Organic Carbon
TRE:	Toxicity Reduction Evaluation
TSD:	Technical Support Document
TSS:	Total Suspended Solids
WER:	Water-Effect Ratio
WET:	Whole Effluent Toxicity
WLA:	Wasteload Allocation

GLOSSARY

Acute-chronic ratio - an appropriate measure of the acute toxicity of a material divided by an appropriate measure of the chronic toxicity of the same material under the same conditions.

Appropriate regulatory authority - Usually the State water pollution control agency, even for States under the National Toxics Rule; if, however, a State were to waive its section 401 authority, the Water Management Division of the EPA Regional Office would become the appropriate regulatory authority.

Clean techniques - a set of procedures designed to prevent contamination of samples so that concentrations of trace metals can be measured accurately and precisely.

Critical species - a species that is commercially or recreationally important at the site, a species that exists at the site and is listed as threatened or endangered under section 4 of the Endangered Species Act, or a species for which there is evidence that the loss of the species from the site is likely to cause an unacceptable impact on a commercially or recreationally important species, a threatened or endangered species, the abundances of a variety of other species, or the structure or function of the community.

Design flow - the flow used for steady-state wasteload allocation modeling.

Dissolved metal - defined here as "metal that passes through either a 0.45- μ m or a 0.40- μ m membrane filter".

Endpoint - the concentration of test material that is expected to cause a specified amount of adverse effect.

Final Water-Effect Ratio - the WER that is used in the calculation of a site-specific aquatic life criterion.

Flow-through test - a test in which test solutions flow into the test chambers either intermittently (every few minutes) or continuously and the excess flows out.

Labile metal - metal that is in water and will readily convert from one form to another when in a nonequilibrium condition.

Particulate metal - metal that is measured by the total recoverable method but not by the dissolved method.

Primary test - the toxicity test used in the determination of a Final Water-Effect Ratio (FWER); the specification of the test includes the test species, the life stage of the species, the duration of the test, and the adverse effect on which the endpoint is based.

Refractory metal - metal that is in water and will not readily convert from one form to another when in a nonequilibrium condition, i.e., metal that is in water and is not labile.

Renewal test - a test in which either the test solution in a test chamber is renewed at least once during the test or the test organisms are transferred into a new test solution of the same composition at least once during the test.

Secondary test - a toxicity test that is usually conducted along with the primary test only once to test the assumptions that, within experimental variation, (a) similar WERs will be obtained using tests that have similar sensitivities to the test material, and (b) tests that are less sensitive to the test material will usually give WERs that are closer to 1.

Simulated downstream water - a site water prepared by mixing effluent and upstream water in a known ratio.

Site-specific aquatic life criterion - a water quality criterion for aquatic life that has been derived to be specifically appropriate to the water quality characteristics and/or species composition at a particular location.

Site water - upstream water, actual downstream water, or simulated downstream water in which a toxicity test is conducted side-by-side with the same toxicity test in a laboratory dilution water to determine a WER.

Static test - a test in which the solution and organisms that are in a test chamber at the beginning of the test remain in the chamber until the end of the test.

Total recoverable metal - metal that is in aqueous solution after the sample is appropriately acidified and digested and insoluble material is separated.

Water-effect ratio - an appropriate measure of the toxicity of a material obtained in a site water divided by the same measure of the toxicity of the same material obtained simultaneously in a laboratory dilution water.

PREFACE

Several issues need consideration when guidance such as this is written:

1. Degrees of importance: Procedures and methods are series of instructions, but some of the instructions are more important than others. Some instructions are so important that, if they are not followed, the results will be questionable or unacceptable; other instructions are less important, but definitely desirable. Possibly the best way to express various degrees of importance is the approach described in several ASTM Standards, such as in section 3.6 of Standard E729 (ASTM 1993a), which is modified here to apply to WERs: The words "must", "should", "may", "can", and "might" have specific meanings in this document. "Must" is used to express an instruction that is to be followed, unless a site-specific consideration requires a deviation, and is used only in connection with instructions that directly relate to the validity of toxicity tests, WERs, FWERs, and the Recalculation Procedure. "Should" is used to state instructions that are recommended and are to be followed if reasonably possible. Deviation from one "should" will not invalidate a WER, but deviation from several probably will. Terms such as "is desirable", "is often desirable", and "might be desirable" are used in connection with less important instructions. "May" is used to mean "is (are) allowed to", "can" is used to mean "is (are) able to", and "might" is used to mean "could possibly". Thus the classic distinction between "may" and "can" is preserved, and "might" is not used as a synonym for either "may" or "can". This does not eliminate all problems concerning the degree of importance, however. For example, a small deviation from a "must" might not invalidate a WER, whereas a large deviation would. (Each "**must**" and "**must not**" is in bold print for convenience, not for emphasis, in this document.)
2. Educational and explanatory material: Many people have asked for much detail in this document to ensure that as many WERs as possible are determined in an acceptable manner. In addition, some people want justifications for each detail. Much of the detail that is desired by some people is based on "best professional judgment", which is rarely considered an acceptable justification by people who disagree with a specified detail. Even if details are taken from an EPA method or an ASTM standard, they were often included in those documents on the basis of best professional judgment. In contrast, some people want detailed methodology presented without explanatory material. It was decided to include as much detail as is feasible, and to provide rationale and explanation for major items.

3. Alternatives: When more than one alternative is both scientifically sound and appropriately protective, it seems reasonable to present the alternatives rather than presenting the one that is considered best. The reader can then select one based on cost-effectiveness, personal preference, details of the particular situation, and perceived advantages and disadvantages.
4. Separation of "science", "best professional judgment" and "regulatory decisions": These can never be completely separated in this kind of document; for example, if data are analyzed for a statistically significant difference, the selection of alpha is an important decision, but a rationale for its selection is rarely presented, probably because the selection is not a scientific decision. In this document, an attempt has been made to focus on good science, best professional judgment, and presentation of the rationale; when possible, these are separated from "regulatory decisions" concerning margin of safety, level of protection, beneficial use, regulatory convenience, and the goal of zero discharge. Some "regulatory decisions" relating to implementation, however, should be integrated with, not separated from, "science" because the two ought to be carefully considered together wherever science has implications for implementation.
5. Best professional judgment: Much of the guidance contained herein is qualitative rather than quantitative, and much judgment will usually be required to derive a site-specific water quality criterion for aquatic life. In addition, although this version of the guidance for determining and using WERs attempts to cover all major questions that have arisen during use of the previous version and during preparation of this version, it undoubtedly does not cover all situations, questions, and extenuating circumstances that might arise in the future. All necessary decisions should be based on both a thorough knowledge of aquatic toxicology and an understanding of this guidance; each decision should be consistent with the spirit of this guidance, which is to make best use of "good science" to derive the most appropriate site-specific criteria. This guidance should be modified whenever sound scientific evidence indicates that a site-specific criterion produced using this guidance will probably substantially underprotect or overprotect the aquatic life at the site of concern. Derivation of site-specific criteria for aquatic life is a complex process and requires knowledge in many areas of aquatic toxicology; any deviation from this guidance should be carefully considered to ensure that it is consistent with other parts of this guidance and with "good science".
6. Personal bias: Bias can never be eliminated, and some decisions are at the fine line between "bias" and "best

professional judgment". The possibility of bias can be eliminated only by adoption of an extreme position such as "no regulation" or "no discharge". One way to deal with bias is to have decisions made by a team of knowledgeable people.

7. Teamwork: The determination of a WER should be a cooperative team effort beginning with the completion of the initial workplan, interpretation of initial data, revision of the workplan, etc. The interaction of a variety of knowledgeable, reasonable people will help obtain the best results for the expenditure of the fewest resources. Members of the team should acknowledge their biases so that the team can make best use of the available information, taking into account its relevancy to the immediate situation and its quality.

INTRODUCTION

National aquatic life criteria for metals are intended to protect the aquatic life in almost all surface waters of the United States (U.S. EPA 1985). This level of protection is accomplished in two ways. First, the national dataset is required to contain aquatic species that have been found to be sensitive to a variety of pollutants. Second, the dilution water and the metal salt used in the toxicity tests are required to have physical and chemical characteristics that ensure that the metal is at least as toxic in the tests as it is in nearly all surface waters. For example, the dilution water is to be low in suspended solids and in organic carbon, and some forms of metal (e.g., insoluble metal and metal bound by organic complexing agents) cannot be used as the test material. (The term "metal" is used herein to include both "metals" and "metalloids".)

Alternatively, a national aquatic life criterion might not adequately protect the aquatic life at some sites. An untested species that is important at a site might be more sensitive than any of the tested species. Also, the metal might be more toxic in site water than in laboratory dilution water because, for example, the site water has a lower pH and/or hardness than most laboratory waters. Thus although a national aquatic life criterion is intended to be lower than necessary for most sites, a national criterion might not adequately protect the aquatic life at some sites.

Because a national aquatic life criterion might be more or less protective than intended for the aquatic life in most bodies of water, the U.S. EPA provided guidance (U.S. EPA 1983a, 1984) concerning three procedures that may be used to derive a site-specific criterion:

1. The Recalculation Procedure is intended to take into account relevant differences between the sensitivities of the aquatic organisms in the national dataset and the sensitivities of organisms that occur at the site.
2. The Indicator Species Procedure provides for the use of a water-effect ratio (WER) that is intended to take into account relevant differences between the toxicity of the metal in laboratory dilution water and in site water.
3. The Resident Species Procedure is intended to take into account both kinds of differences simultaneously.

A site-specific criterion is intended to come closer than the national criterion to providing the intended level of protection to the aquatic life at the site, usually by taking into account the biological and/or chemical conditions (i.e., the species composition and/or water quality characteristics) at the site. The fact that the U.S. EPA has made these procedures available should not be interpreted as implying that the agency advocates that states derive site-specific criteria before setting state

standards. Also, derivation of a site-specific criterion does not change the intended level of protection of the aquatic life at the site. Because a WER is expected to appropriately take into account (a) the site-specific toxicity of the metal, and (b) synergism, antagonism, and additivity with other constituents of the site water, using a WER is more likely to provide the intended level of protection than not using a WER.

Although guidance concerning site-specific criteria has been available since 1983 (U.S. EPA 1983a, 1984), interest has increased in recent years as states have devoted more attention to chemical-specific water quality criteria for aquatic life. In addition, interest in water-effect ratios (WERs) increased when the "Interim Guidance" concerning metals (U.S. EPA 1992) made a fundamental change in the way that WERs are experimentally determined (see Appendix A), because the change is expected to substantially increase the magnitude of many WERs. Interest was further focused on WERs when they were integrated into some of the aquatic life criteria for metals that were promulgated by the National Toxics Rule (57 FR 60848, December 22, 1992). The newest guidance issued by the U.S. EPA (Prothro 1993) concerning aquatic life criteria for metals affected the determination and use of WERs only insofar as it affected the use of total recoverable and dissolved criteria.

The early guidance concerning WERs (U.S. EPA 1983a, 1984) contained few details and needs revision, especially to take into account newer guidance concerning metals (U.S. EPA 1992; Prothro 1993). The guidance presented herein supersedes all guidance concerning WERs and the Indicator Species Procedure given in Chapter 4 of the Water Quality Standards Handbook (U.S. EPA 1983a) and in U.S. EPA (1984). All guidance presented in U.S. EPA (1992) is superseded by that presented by Prothro (1993) and by this document. Metals are specifically addressed herein because of the National Toxics Rule (NTR) and because of current interest in aquatic life criteria for metals; although most of this guidance also applies to other pollutants, some obviously applies only to metals.

Even though this document was prepared mainly because of the NTR, the guidance contained herein concerning WERs is likely to have impact beyond its use with the NTR. Therefore, it is appropriate to also present new guidance concerning the Recalculation Procedure (see Appendix B) because the previous guidance (U.S. EPA 1983a, 1984) concerning this procedure also contained few details and needs revision. The NTR does not allow use of the Recalculation Procedure in jurisdictions subject to the NTR.

The previous guidance concerning site-specific procedures did not allow the Recalculation Procedure and the WER procedure to be used together in the derivation of a site-specific aquatic life criterion; the only way to take into account both species

composition and water quality characteristics in the determination of a site-specific criterion was to use the Resident Species Procedure. A specific change contained herein is that, except in jurisdictions that are subject to the NTR, the Recalculation Procedure and the WER Procedure may now be used together. Additional reasons for addressing both the Recalculation Procedure and the WER Procedure in this document are that both procedures are based directly on the guidelines for deriving national aquatic life criteria (U.S. EPA 1985) and, when the two are used together, use of the Recalculation Procedure has specific implications concerning the determination of the WER.

This guidance is intended to produce WERs that may be used to derive site-specific aquatic life criteria for metals from most national and state aquatic life criteria that were derived from laboratory toxicity data. Except in jurisdictions that are subject to the NTR, the WERs may also be used with site-specific aquatic life criteria that are derived for metals using the Recalculation Procedure described in Appendix B. WERs obtained using the methods described herein should not be used to adjust aquatic life criteria that were derived for metals in other ways. For example, because they are designed to be applied to criteria derived on the basis of laboratory toxicity tests, WERs determined using the methods described herein cannot be used to adjust the residue-based mercury Criterion Continuous Concentration (CCC) or the field-based selenium freshwater criterion. For the purposes of the NTR, WERs may be used with the aquatic life criteria for arsenic, cadmium, chromium(III), chromium(VI), copper, lead, nickel, silver, and zinc and with the Criterion Maximum Concentration (CMC) for mercury. WERs may also be used with saltwater criteria for selenium.

The concept of a WER is rather simple:

Two side-by-side toxicity tests are conducted - one test using laboratory dilution water and the other using site water. The endpoint obtained using site water is divided by the endpoint obtained using laboratory dilution water. The quotient is the WER, which is multiplied times the national, state, or recalculated aquatic life criterion to calculate the site-specific criterion.

Although the concept is simple, the determination and use of WERs involves many considerations.

The primary purposes of this document are to:

1. Identify steps that should be taken before the determination of a WER is begun.
2. Describe the methods recommended by the U.S. EPA for the determination of WERs.
3. Address some issues concerning the use of WERs.
4. Present new guidance concerning the Recalculation Procedure.

Before Determining a WER

Because a national criterion is intended to protect aquatic life in almost all bodies of water and because a WER is intended to account for a difference between the toxicity of a metal in a laboratory dilution water and its toxicity in a site water, dischargers who want higher permit limits than those derived on the basis of an existing aquatic life criterion will probably consider determining a WER. Use of a WER should be considered only as a last resort for at least three reasons:

- a. Even though some WERs will be substantially greater than 1.0, some will be about 1.0 and some will be less than 1.0.
- b. The determination of a WER requires substantial resources.
- c. There are other things that a discharger can do that might be more cost-effective than determining a WER.

The two situations in which the determination of a WER might appear attractive to dischargers are when (a) a discharge appears to exceed existing or proposed water quality-based permit limits, and (b) an instream concentration appears to exceed an existing or proposed aquatic life criterion. Such situations result from measurement of the concentration of a metal in an effluent or a surface water. It would therefore seem reasonable to ensure that such measurements were not subject to contamination. Usually it is much easier to verify chemical measurements by using "clean techniques" for collecting, handling, storing, preparing, and analyzing samples, than to determine a WER. Clean techniques and some related QA/QC considerations are discussed in Appendix C.

In addition to investigating the use of "clean techniques", other steps that a discharger should take prior to beginning the experimental determination of a WER include:

1. Evaluate the potential for reducing the discharge of the metal.
2. Investigate such possible constraints on permit limits as antibacksliding and antidegradation requirements and human health and wildlife criteria.
3. Obtain assistance from an aquatic toxicologist who understands the basics of WERs (see Appendix D), the U.S. EPA's national aquatic life guidelines (U.S. EPA 1985), the guidance presented by Prothro (1993), the national criteria document for the metal(s) of concern (see Appendix E), the procedures described by the U.S. EPA (1993a,b,c) for acute and chronic toxicity tests on effluents and surface waters, and the procedures described by ASTM (1993a,b,c,d,e) for acute and chronic toxicity tests in laboratory dilution water.
4. Develop an initial definition of the site to which the site-specific criterion is to apply.
5. Consider use of the Recalculation Procedure (see Appendix B).
6. Evaluate the cost-effectiveness of the determination of a WER. Comparative toxicity tests provide the most useful data, but chemical analysis of the downstream water might be helpful

because the following are often true for some metals:

- a. The lower the percent of the total recoverable metal in the downstream water that is dissolved, the higher the WER.
- b. The higher the concentration of total organic carbon (TOC) and/or total suspended solids (TSS), the higher the WER.

It is also true that the higher the concentration of nontoxic dissolved metal, the higher the WER. Although some chemical analyses might provide useful information concerning the toxicities of some metals in water, at the present only toxicity tests can accurately reflect the toxicities of different forms of a metal (see Appendix D).

7. Submit a workplan for the experimental determination of the WER to the appropriate regulatory authority (and possibly to the Water Management Division of the EPA Regional Office) for comment. The workplan should include detailed descriptions of the site; existing criterion and standard; design flows; site water; effluent; sampling plan; procedures that will be used for collecting, handling, and analyzing samples of site water and effluent; primary and secondary toxicity tests; quality assurance/quality control (QA/QC) procedures; Standard Operating Procedures (SOPs); and data interpretation.

After the workplan is completed, the initial phase should be implemented; then the data obtained should be evaluated, and the workplan should be revised if appropriate. Developing and modifying the workplan and analyzing and interpreting the data should be a cooperative effort by a team of knowledgeable people.

Two Kinds of WERS

Most aquatic life criteria contain both a CMC and a CCC, and it is usually possible to determine both a cmcWER and a cccWER. The two WERS cannot be assumed to be equal because the magnitude of a WER will probably depend on the sensitivity of the toxicity test used and on the percent effluent in the site water (see Appendix D), both of which can depend on which WER is to be determined. In some cases, it is expected that a larger WER can be applied to the CCC than to the CMC, and so it would be environmentally conservative to apply cmcWERS to CCCs. In such cases it is possible to determine a cmcWER and apply it to both the CMC and the CCC in order to derive a site-specific CMC, a site-specific CCC, and new permit limits. If these new permit limits are controlled by the new site-specific CCC, a cccWER could be determined using a more sensitive test, possibly raising the site-specific CCC and the permit limits again. A cccWER may, of course, be determined whenever desired. Unless the experimental variation is increased, use of a cccWER will usually improve the accuracy of the resulting site-specific CCC.

In some cases, a larger WER cannot be applied to the CCC than to the CMC and so it might not be environmentally conservative to apply a cmcWER to a CCC (see section A.4 of Method 1).

Steady-state and Dynamic Models

Some of the guidance contained herein specifically applies to situations in which the permit limits were calculated using steady-state modeling; in particular, some samples are to be obtained when the actual stream flow is close to the design flow. If permit limits were calculated using dynamic modeling, the guidance will have to be modified, but it is unclear at present what modifications are most appropriate. For example, it might be useful to determine whether the magnitude of the WER is related to the flow of the upstream water and/or the effluent.

Two Methods

Two methods are used to determine WERs. Method 1 will probably be used to determine all cmcWERS and most cccWERS because it can be applied to situations that are in the vicinities of plumes. Because WERs are likely to depend on the concentration of effluent in the water and because the percent effluent in a water sample obtained in the immediate vicinity of a plume is unknown, simulated downstream water is used so that the percent effluent in the sample is known. For example, if a sample that was supposed to represent a complete-mix situation was accidentally taken in the plume upstream of complete mix, the sample would probably have a higher percent effluent and a higher WER than a sample taken downstream of complete mix; use of the higher WER to derive a site-specific criterion for the complete-mix situation would result in underprotection. If the sample were accidentally taken upstream of complete mix but outside the plume, overprotection would probably result.

Method 1 will probably be used to determine all cmcWERS and most cccWERS in flowing fresh waters, such as rivers and streams. Method 1 is intended to apply not only to ordinary rivers and streams but also to streams that some people might consider extraordinary, such as streams whose design flows are zero and streams that some state and/or federal agencies refer to as "effluent-dependent", "habitat-creating", or "effluent-dominated". Method 1 is also used to determine cmcWERS in such large sites as oceans and large lakes, reservoirs, and estuaries (see Appendix F).

Method 2 is used to determine WERs that apply outside the area of plumes in large bodies of water. Such WERs will be cccWERS and will be determined using samples of actual site water obtained at various times, locations, and depths in order to identify the range of WERs that apply to the body of water. These experimentally determined WERs are then used to decide how many site-specific criteria should be derived for the body of water and what the criterion (or criteria) should be. Method 2 requires substantially more resources than Method 1.

The complexity of each method increases when the number of metals and/or the number of discharges is two or more:

- a. The simplest situation is when a WER is to be determined for only one metal and only one discharge has permit limits for that metal. (This is the single-metal single-discharge situation.)
- b. A more complex situation is when a WER is to be determined for only one metal, but more than one discharge has permit limits for that metal. (This is the single-metal multiple-discharge situation.)
- c. An even more complex situation is when WERs are to be determined for more than one metal, but only one discharge has permit limits for any of the metals. (This is the multiple-metal single-discharge situation.)
- d. The most complex situation is when WERs are to be determined for more than one metal and more than one discharge has permit limits for some or all of the metals. (This is the multiple-metal multiple-discharge situation.)

WERs need to be determined for each metal at each site because extrapolation of a WER from one metal to another, one effluent to another, or one surface water to another is too uncertain.

Both methods work well in multiple-metal situations, but special tests or additional tests will be necessary to show that the resulting combination of site-specific criteria will not be too toxic. Method 2 is better suited to multiple-discharge situations than is Method 1. Appendix F provides additional guidance concerning multiple-metal and multiple-discharge situations, but it does not discuss allocation of waste loads, which is performed when a wasteload allocation (WLA) or a total maximum daily load (TMDL) is developed (U.S. EPA 1991a).

Two Analytical Measurements

A total recoverable WER can be determined if the metal in both of the side-by-side toxicity tests is analyzed using the total recoverable measurement; similarly, a dissolved WER can be determined if the metal in both tests is analyzed using the dissolved measurement. A total recoverable WER is used to calculate a total recoverable site-specific criterion from an aquatic life criterion that is expressed using the total recoverable measurement, whereas a dissolved WER is used to calculate a dissolved site-specific criterion from a criterion that is expressed in terms of the dissolved measurement. Figure 1 illustrates the relationships between total recoverable and dissolved criteria, WERs, and the Recalculation Procedure.

Both Method 1 and Method 2 can be used to determine a total recoverable WER and/or a dissolved WER. The only difference in the experimental procedure is whether the WER is based on measurements of total recoverable metal or dissolved metal in the

test solutions. Both total recoverable and dissolved measurements are to be performed for all tests to help judge the quality of the tests, to provide a check on the analytical chemistry, and to help understand the results; performing both measurements also increases the alternatives available for use of the results. For example, a dissolved WER that is not useful with a total recoverable criterion might be useful in the future if a dissolved criterion becomes available. Also, as explained in Appendix D, except for experimental variation, use of a total recoverable WER with a total recoverable criterion should produce the same total recoverable permit limits as use of a dissolved WER with a dissolved criterion; the internal consistency of the approaches and the data can be evaluated if both total recoverable and dissolved criteria and WERs are determined. It is expected that in many situations total recoverable WERs will be larger and more variable than dissolved WERs.

The Quality of the Toxicity Tests

Traditionally, for practical reasons, the requirements concerning such aspects as acclimation of test organisms to test temperature and dilution water have not been as stringent for toxicity tests on surface waters and effluents as for tests using laboratory dilution water. Because a WER is a ratio calculated from the results of side-by-side tests, it might seem that acclimation is not important for a WER as long as the organisms and conditions are identical in the two tests. Because WERs are used to adjust aquatic life criteria that are derived from results of laboratory tests, the tests conducted in laboratory dilution water for the determination of WERs should be conducted in the same way as the laboratory toxicity tests used in the derivation of aquatic life criteria. In the WER process, the tests in laboratory dilution water provide the vital link between national criteria and site-specific criteria, and so it is important to compare at least some results obtained in the laboratory dilution water with results obtained in at least one other laboratory.

Three important principles for making decisions concerning the methodology for the side-by-side tests are:

1. The tests using laboratory dilution water should be conducted so that the results would be acceptable for use in the derivation of national criteria.
2. As much as is feasible, the tests using site water should be conducted using the same procedures as the tests using the laboratory dilution water.
3. All tests should follow any special requirements that are necessary because the results are to be used to calculate a WER. Some such special requirements are imposed because the criterion for a rather complex situation is being changed based on few data, so more assurance is required that the data are high quality.

The most important special requirement is that the concentrations of the metal are to be measured using both the total recoverable and dissolved methods in all toxicity tests used for the determination of a WER. This requirement is necessary because half of the tests conducted for the determination of WERs use a site water in which the concentration of metal probably is not negligible. Because it is likely that the concentration of metal in the laboratory dilution water is negligible, assuming that the concentration in both waters is negligible and basing WERs on the amount of metal added would produce an unnecessarily low value for the WER. In addition, WERs are based on too few data to assume that nominal concentrations are accurate. Nominal concentrations obviously cannot be used if a dissolved WER is to be determined. Measured dissolved concentrations at the beginning and end of the test are used to judge the acceptability of the test, and it is certainly reasonable to measure the total recoverable concentration when the dissolved concentration is measured. Further, measuring the concentrations might lead to an interpretation of the results that allows a substantially better use of the WERs.

Conditions for Determining a WER

The appropriate regulatory authority might recommend that one or more conditions be met when a WER is determined in order to reduce the possibility of having to determine a new WER later:

1. Requirements that are in the existing permit concerning WET testing, Toxicity Identification Evaluation (TIE), and/or Toxicity Reduction Evaluation (TRE) (U.S. EPA 1991a).
2. Implementation of pollution prevention efforts, such as pretreatment, waste minimization, and source reduction.
3. A demonstration that applicable technology-based requirements are being met.

If one or more of these is not satisfied when the WER is determined and is implemented later, it is likely that a new WER will have to be determined because of the possibility of a change in the composition of the effluent.

Even if all recommended conditions are satisfied, determination of a WER might not be possible if the effluent, upstream water, and/or downstream water are toxic to the test organisms. In some such cases, it might be possible to determine a WER, but remediation of the toxicity is likely to be required anyway. It is unlikely that a WER determined before remediation would be considered acceptable for use after remediation. If it is desired to determine a WER before remediation and the toxicity is in the upstream water, it might be possible to use a laboratory dilution water or a water from a clean tributary in place of the upstream water; if a substitute water is used, its water quality characteristics should be similar to those of the upstream water (i.e., the pH should be within 0.2 pH units and the hardness,

alkalinity, and concentrations of TSS and TOC should be within 10 % or 5 mg/L, whichever is greater, of those in the upstream water). If the upstream water is chronically toxic, but not acutely toxic, it might be possible to determine a cmcWER even if a cccWER cannot be determined; a cmcWER might not be useful, however, if the permit limits are controlled by the CCC; in such a case, it would probably not be acceptable to assume that the cmcWER is an environmentally conservative estimate of the cccWER. If the WER is determined using downstream water and the toxicity is due to the effluent, tests at lower concentrations of the effluent might give an indication of the amount of remediation needed.

Conditions for Using a WER

Besides requiring that the WER be valid, the appropriate regulatory authority might consider imposing other conditions for the approval of a site-specific criterion based on the WER:

1. Periodic reevaluation of the WER.

- a. WERs determined in upstream water take into account constituents contributed by point and nonpoint sources and natural runoff; thus a WER should be reevaluated whenever newly implemented controls or other changes substantially affect such factors as hardness, alkalinity, pH, suspended solids, organic carbon, or other toxic materials.
- b. Most WERs determined using downstream water are influenced more by the effluent than the upstream water. Downstream WERs should be reevaluated whenever newly implemented controls or other changes might substantially impact the effluent, i.e., might impact the forms and concentrations of the metal, hardness, alkalinity, pH, suspended solids, organic carbon, or other toxic materials. A special concern is the possibility of a shift from discharge of nontoxic metal to discharge of toxic metal such that the concentration of the metal does not increase; analytical chemistry might not detect the change but toxicity tests would.

Even if no changes are known to have occurred, WERs should be reevaluated periodically. (The NTR recommends that NPDES permits include periodic determinations of WERs in the monitoring requirements.) With advance planning, it should usually be possible to perform such reevaluations under conditions that are at least reasonably similar to those that control the permit limits (e.g., either design-flow or high-flow conditions) because there should be a reasonably long period of time during which the reevaluation can be performed. Periodic determination of WERs should be designed to answer questions, not just generate data.

2. Increased chemical monitoring of the upstream water, effluent, and/or downstream water, as appropriate, for water quality characteristics that probably affect the toxicity of the metal

(e.g., hardness, alkalinity, pH, TOC, and TSS) to determine whether conditions change. The conditions at the times the samples were obtained should be kept on record for reference. The WER should be reevaluated whenever hardness, alkalinity, pH, TOC, and/or TSS decrease below the values that existed when the WERs were determined.

3. Periodic reevaluation of the environmental fate of the metal in the effluent (see Appendix A).
4. WET testing.
5. Instream bioassessments.

Decisions concerning the possible imposition of such conditions should take into account:

- a. The ratio of the new and old criteria. The greater the increase in the criterion, the more concern there should be about (1) the fate of any nontoxic metal that contributes to the WER and (2) changes in water quality that might occur within the site. The imposition of one or more conditions should be considered if the WER is used to raise the criterion by, for example, a factor of two, and especially if it is raised by a factor of five or more. The significance of the magnitude of the ratio can be judged by comparison with the acute-chronic ratio, the factor of two that is the ratio of the FAV to the CMC, and the range of sensitivities of species in the criteria document for the metal (see Appendix E).
- b. The size of the site.
- c. The size of the discharge.
- d. The rate of downstream dilution.
- e. Whether the CMC or the CCC controls the permit limits.

When WERs are determined using upstream water, conditions on the use of a WER are more likely when the water contains an effluent that increases the WER by adding TOC and/or TSS, because the WER will be larger and any decrease in the discharge of such TOC and/or TSS might decrease the WER and result in underprotection. A WER determined using downstream water is likely to be larger and quite dependent on the composition of the effluent; there should be concern about whether a change in the effluent might result in underprotection at some time in the future.

Implementation Considerations

In some situations a discharger might not want to or might not be allowed to raise a criterion as much as could be justified by a WER:

1. The maximum possible increase is not needed and raising the criterion more than needed might greatly raise the cost if a greater increase would require more tests and/or increase the conditions imposed on approval of the site-specific criterion.
2. Such other constraints as antibacksliding or antidegradation requirements or human health or wildlife criteria might limit the amount of increase regardless of the magnitude of the WER.

3. The permit limits might be limited by an aquatic life criterion that applies outside the site. It is EPA policy that permit limits cannot be so high that they inadequately protect a portion of the same or a different body of water that is outside the site; nothing contained herein changes this policy in any way.

If no increase in the existing discharge is allowed, the only use of a WER will be to determine whether an existing discharge needs to be reduced. Thus a major use of WERs might be where technology-based controls allow concentrations in surface waters to exceed national, state, or recalculated aquatic life criteria. In this case, it might only be necessary to determine that the WER is greater than a particular value; it might not be necessary to quantify the WER. When possible, it might be desirable to show that the maximum WER is greater than the WER that will be used in order to demonstrate that a margin of safety exists, but again it might not be necessary to quantify the maximum WER.

In jurisdictions not subject to the NTR, WERs should be used to derive site-specific criteria, not just to calculate permit limits, because data obtained from ambient monitoring should be interpreted by comparison with ambient criteria. (This is not a problem in jurisdictions subject to the NTR because the NTR defines the ambient criterion as "WER x the EPA criterion".) If a WER is used to adjust permit limits without adjusting the criterion, the permit limits would allow the criterion to be exceeded. Thus the WER should be used to calculate a site-specific criterion, which should then be used to calculate permit limits. In some states, site-specific criteria can only be adopted as revised criteria in a separate, independent water quality standards review process. In other states, site-specific criteria can be developed in conjunction with the NPDES permitting process, as long as the adoption of a site-specific criterion satisfies the pertinent water quality standards procedural requirements (i.e., a public notice and a public hearing). In either case, site-specific criteria are to be adopted prior to NPDES permit issuance. Moreover, the EPA Regional Administrator has authority to approve or disapprove all new and revised site-specific criteria and to review NPDES permits to verify compliance with the applicable water quality criteria.

Other aspects of the use of WERs in connection with permit limits, WLAs, and TMDLs are outside the scope of this document. The Technical Support Document (U.S. EPA 1991a) and Prothro (1993) provide more information concerning implementation procedures. Nothing contained herein should be interpreted as changing the three-part approach that EPA uses to protect aquatic life: (1) numeric chemical-specific water quality criteria for individual pollutants, (2) whole effluent toxicity (WET) testing, and (3) instream bioassessments.

Even though there are similarities between WET testing and the determination of WERs, there are important differences. For example, WERs can be used to derive site-specific criteria for individual pollutants, but WET testing cannot. The difference between WET testing and the determination of WERs is less when the toxicity tests used in the determination of the WER are ones that are used in WET testing. If a WER is used to make a large change in a criterion, additional WET testing and/or instream bioassessments are likely to be recommended.

The Sample-Specific WER Approach

A major problem with the determination and use of aquatic life criteria for metals is that no analytical measurement or combination of measurements has yet been shown to explain the toxicity of a metal to aquatic plants, invertebrates, amphibians, and fishes over the relevant range of conditions in surface waters (see Appendix D). It is not just that insufficient data exist to justify a relationship; rather, existing data possibly contradict some ideas that could possibly be very useful if true. For example, the concentration of free metal ion could possibly be a useful basis for expressing water quality criteria for metals if it could be feasible and could be used in a way that does not result in widespread underprotection of aquatic life. Some available data, however, might contradict the idea that the toxicity of copper to aquatic organisms is proportional to the concentration or the activity of the cupric ion. Evaluating the usefulness of any approach based on metal speciation is difficult until it is known how many of the species of the metal are toxic, what the relative toxicities are, whether they are additive (if more than one is toxic), and the quantitative effects of the factors that have major impacts on the bioavailability and/or toxicity of the toxic species. Just as it is not easy to find a useful quantitative relationship between the analytical chemistry of metals and the toxicity of metals to aquatic life, it is also not easy to find a qualitative relationship that can be used to provide adequate protection for the aquatic life in almost all bodies of water without providing as much overprotection for some bodies of water as results from use of the total recoverable and dissolved measurements.

The U.S. EPA cannot ignore the existence of pollution problems and delay setting aquatic life criteria until all scientific issues have been adequately resolved. In light of uncertainty, the agency needs to derive criteria that are environmentally conservative in most bodies of water. Because of uncertainty concerning the relationship between the analytical chemistry and the toxicity of metals, aquatic life criteria for metals are expressed in terms of analytical measurements that result in the criteria providing more protection than necessary for the aquatic life in most bodies of water. The agency has provided for the

use of WERs to address the general conservatism, but expects that some WERs will be less than 1.0 because national, state, and recalculated criteria are not necessarily environmentally conservative for all bodies of water.

It has become obvious, however, that the determination and use of WERs is not a simple solution to the existing general conservatism. It is likely that a permanent solution will have to be based on an adequate quantitative explanation of how metals and aquatic organisms interact. In the meantime, the use of total recoverable and dissolved measurements to express criteria and the use of site-specific criteria are intended to provide adequate protection for almost all bodies of water without excessive overprotection for too many bodies of water. Work needs to continue on the permanent solution and, just in case, on improved alternative approaches.

Use of WERs to derive site-specific criteria is intended to allow a reduction or elimination of the general overprotection associated with application of a national criterion to individual bodies of water, but a major problem is that a WER will rarely be constant over time, location, and depth in a body of water due to plumes, mixing, and resuspension. It is possible that dissolved concentrations and WERs will be less variable than total recoverable ones. It might also be possible to reduce the impact of the heterogeneity if WERs are additive across time, location, and depth (see Appendix G). Regardless of what approaches, tools, hypotheses, and assumptions are utilized, variation will exist and WERs will have to be used in a conservative manner. Because of variation between bodies of water, national criteria are derived to be environmentally conservative for most bodies of water, whereas the WER procedure, which is intended to reduce the general conservatism of national criteria, has to be conservative because of variation among WERs within a body of water.

The conservatism introduced by variation among WERs is due not to the concept of WERs, but to the way they are used. The reason that national criteria are conservative in the first place is the uncertainty concerning the linkage of analytical chemistry and toxicity; the toxicity of solutions can be measured, but toxicity cannot be modelled adequately using available chemical measurements. Similarly, the current way that WERs are used depends on a linkage between analytical chemistry and toxicity because WERs are used to derive site-specific criteria that are expressed in terms of chemical measurements.

Without changing the amount or kind of toxicity testing that is performed when WERs are determined using Method 2, a different way of using the WERs could avoid some of the problems introduced by the dependence on analytical chemistry. The "sample-specific WER approach" could consist of sampling a body of water at a number of locations, determining the WER for each sample, and

measuring the concentration of the metal in each sample. Then for each individual sample, a quotient would be calculated by dividing the concentration of metal in the sample by the product of the national criterion times the WER obtained for that sample. Except for experimental variation, when the quotient for a sample is less than 1, the concentration of metal in that sample is acceptable; when the quotient for a sample is greater than 1, the concentration of metal in that sample is too high. As a check, both the total recoverable measurement and the dissolved measurement should be used because they should provide the same answer if everything is done correctly and accurately. This approach can also be used whenever Method 1 is used; although Method 1 is used with simulated downstream water, the sample-specific WER approach can be used with either simulated downstream water or actual downstream water.

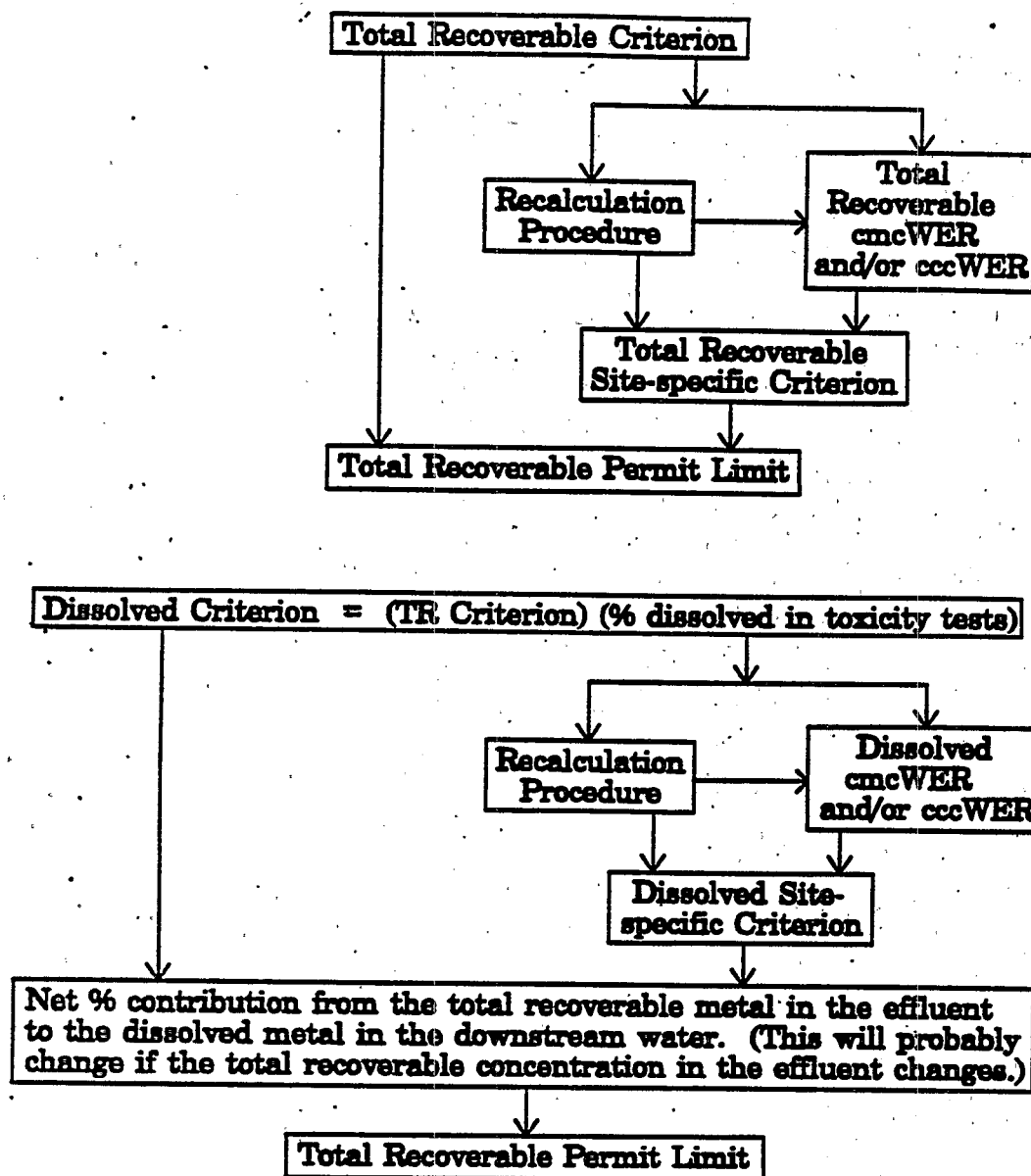
This sample-specific WER approach has several interesting features:

1. It is not a different way of determining WERs; it is merely a different way of using the WERs that are determined.
2. Variation among WERs within a body of water is not a problem.
3. It eliminates problems concerning the unknown relationship between toxicity and analytical chemistry.
4. It works equally well in areas that are in or near plumes and in areas that are away from plumes.
5. It works equally well in single-discharge and multiple-discharge situations.
6. It automatically accounts for synergism, antagonism, and additivity between toxicants.

This way of using WERs is equivalent to expressing the national criterion for a pollutant in terms of toxicity tests whose endpoints equal the CMC and the CCC; if the site water causes less adverse effect than is defined to be the endpoint, the concentration of that pollutant in the site water does not exceed the national criterion. This sample-specific WER approach does not directly fit into the current framework wherein criteria are derived and then permit limits are calculated from the criteria.

If the sample-specific WER approach were to produce a number of quotients that are greater than 1, it would seem that the concentration of metal in the discharge(s) should be reduced enough that the quotient is not greater than 1. Although this might sound straightforward, the discharger(s) would find that a substantial reduction in the discharge of a metal would not achieve the intended result if the reduction was due to removal of nontoxic metal. A chemical monitoring approach that cannot differentiate between toxic and nontoxic metal would not detect that only nontoxic metal had been removed, but the sample-specific WER approach would.

Figure 1: Four Ways to Derive a Permit Limit



For both the total recoverable and dissolved measurements, derivation of an optional site-specific criterion is described on the right. If both the Recalculation Procedure and the WER procedure are used, the Recalculation Procedure must be performed first. (The Recalculation Procedure cannot be used in jurisdictions that are subject to the National Toxics Rule.)

METHOD 1: DETERMINING WERS FOR AREAS IN OR NEAR PLUMES

Method 1 is based on the determination of WERS using simulated downstream water and so it can be used to determine a WER that applies in the vicinity of a plume. Use of simulated downstream water ensures that the concentration of effluent in the site water is known, which is important because the magnitude of the WER will often depend on the concentration of effluent in the downstream water. Knowing the concentration of effluent makes it possible to quantitatively relate the WER to the effluent.

Method 1 can be used to determine either cmcWERS or cccWERS or both in single-metal, flowing freshwater situations, including streams whose design flow is zero and "effluent-dependent" streams (see Appendix F). As is also explained in Appendix F, Method 1 is used when cmcWERS are determined for "large sites", although Method 2 is used when cccWERS are determined for "large sites". In addition, Appendix F addresses special considerations regarding multiple-metal and/or multiple-discharge situations.

Neither Method 1 nor Method 2 covers all important methodological details for conducting the side-by-side toxicity tests that are necessary in order to determine a WER. Many references are made to information published by the U.S. EPA (1993a,b,c) concerning toxicity tests on effluents and surface waters and by ASTM (1993a,b,c,d,e,f) concerning tests in laboratory dilution water. Method 1 addresses aspects of toxicity tests that (a) need special attention when determining WERS and/or (b) are usually different for tests conducted on effluents and tests conducted in laboratory dilution water. Appendix H provides additional information concerning toxicity tests with saltwater species.

A. Experimental Design

Because of the variety of considerations that have important implications for the determination of a WER, decisions concerning experimental design should be given careful attention and need to answer the following questions:

1. Should WERS be determined using upstream water, actual downstream water, and/or simulated downstream water?
2. Should WERS be determined when the stream flow is equal to, higher than, and/or lower than the design flow?
3. Which toxicity tests should be used?
4. Should a cmcWER or a cccWER or both be determined?
5. How should a FWER be derived?
6. For metals whose criteria are hardness-dependent, at what hardness should WERS be determined?

The answers to these questions should be based on the reason that WERS are determined, but the decisions should also take into account some practical considerations.

1. Should WERs be determined using upstream water, actual downstream water, and/or simulated downstream water?
 - a. Upstream water provides the least complicated way of determining and using WERs because plumes, mixing zones, and effluent variability do not have to be taken into account. Use of upstream water provides the least useful WERs because it does not take into account the presence of the effluent, which is the source of the metal. It is easy to assume that upstream water will give smaller WERs than downstream water, but in some cases downstream water might give smaller WERs (see Appendix G). Regardless of whether upstream water gives smaller or larger WERs, a WER should be determined using the water to which the site-specific criterion is to apply (see Appendix A).
 - b. Actual downstream water might seem to be the most pertinent water to use when WERs are determined, but whether this is true depends on what use is to be made of the WERs. WERs determined using actual downstream water can be quantitatively interpreted using the sample-specific WER approach described at the end of the Introduction. If, however, it is desired to understand the quantitative implications of a WER for an effluent of concern, use of actual downstream water is problematic because the concentration of effluent in the water can only be known approximately.

Sampling actual downstream water in areas that are in or near plumes is especially difficult. The WER obtained is likely to depend on where the sample is taken because the WER will probably depend on the percent effluent in the sample (see Appendix D). The sample could be taken at the end of the pipe, at the edge of the acute mixing zone, at the edge of the chronic mixing zone, or in a completely mixed situation. If the sample is taken at the edge of a mixing zone, the composition of the sample will probably differ from one point to another along the edge of the mixing zone.

If samples of actual downstream water are to be taken close to a discharge, the mixing patterns and plumes should be well known. Dye dispersion studies (Kilpatrick 1992) are commonly used to determine isopleths of effluent concentration and complete mix; dilution models (U.S. EPA 1993d) might also be helpful when selecting sampling locations. The most useful samples of actual downstream water are probably those taken just downstream of the point at which complete mix occurs or at the most distant point that is within

the site to which the site-specific criterion is to apply. When samples are collected from a complete-mix situation, it might be appropriate to composite samples taken over a cross section of the stream. Regardless of where it is decided conceptually that a sample should be taken, it might be difficult to identify where the point exists in the stream and how it changes with flow and over time. In addition, if it is not known exactly what the sample actually represents, there is no way to know how reproducible the sample is. These problems make it difficult to relate WERs determined in actual downstream water to an effluent of concern because the concentration of effluent in the sample is not known; this is not a problem, however, if the sample-specific WER approach is used to interpret the results.

- c. Simulated downstream water would seem to be the most unnatural of the three kinds of water, but it offers several important advantages because effluent and upstream water are mixed at a known ratio. This is important because the magnitude of the WER will often depend on the concentration of effluent in the downstream water. Mixtures can be prepared to simulate the ratio of effluent and upstream water that exists at the edge of the acute mixing zone, at the edge of the chronic mixing zone, at complete mix, or at any other point of interest. If desired, a sample of effluent can be mixed with a sample on upstream water in different ratios to simulate different points in a stream. Also, the ratio used can be one that simulates conditions at design flow or at any other flow.

The sample-specific WER approach can be used with both actual and simulated downstream water. Additional quantitative uses can be made of WERs determined using simulated downstream water because the percent effluent in the water is known, which allows quantitative extrapolations to the effluent. In addition, simulated downstream water can be used to determine the variation in the WER that is due to variation in the effluent. It also allows comparison of two or more effluents and determination of the interactions of two or more effluents. Additivity of WERs can be studied using simulated downstream water (see Appendix G); studies of toxicity within plumes and studies of whether increased flow of upstream water can increase toxicity are both studies of additivity of WERs. Use of simulated downstream water also makes it possible to conduct controlled studies of changes in WERs due to aging and changes in pH.

In Method 1, therefore, WERs are determined using simulated downstream water that is prepared by mixing samples of effluent and upstream water in an appropriate ratio. Most importantly, Method 1 can be used to determine a WER that applies in the vicinity of a plume and can be quantitatively extrapolated to the effluent.

2. Should WERs be determined when the stream flow is equal to, higher than, and/or lower than the design flow?

WERs are used in the derivation of site-specific criteria when it is desired that permit limits be based on a criterion that takes into account the characteristics of the water and/or the metal at the site. In most cases, permit limits are calculated using steady-state models and are based on a design flow. It is therefore important that WERs be adequately protective under design-flow conditions, which might be expected to require that some sets of samples of effluent and upstream water be obtained when the actual stream flow is close to the design flow. Collecting samples when the stream flow is close to the design flow will limit a WER determination to the low-flow season (e.g., from mid-July to mid-October in some places) and to years in which the flow is sufficiently low.

It is also important, however, that WERs that are applied at design flow provide adequate protection at higher flows. Generalizations concerning the impact of higher flows on WERs are difficult because such flows might (a) reduce hardness, alkalinity, and pH, (b) increase or decrease the concentrations of TOC and TSS, (c) resuspend toxic and/or nontoxic metal from the sediment, and (d) wash additional pollutants into the water. Acidic snowmelt, for example, might lower the WER both by diluting the WER and by reducing the hardness, alkalinity, and pH; if substantial labile metal is present, the WER might be lowered more than the concentration of the metal, possibly resulting in increased toxicity at flows higher than design flow. Samples taken at higher flows might give smaller WERs because the concentration of the effluent is more dilute; however, total recoverable WERs might be larger if the sample is taken just after an event that greatly increases the concentration of TSS and/or TOC because this might increase both (1) the concentration of nontoxic particulate metal in the water and (2) the capacity of the water to sorb and detoxify metal.

WERs are not of concern when the stream flow is lower than the design flow because these are acknowledged times of reduced protection. Reduced protection might not occur, however, if the WER is sufficiently high when the flow is lower than design flow.

3. Which toxicity tests should be used?

- a. As explained in Appendix D, the magnitude of an experimentally determined WER is likely to depend on the sensitivity of the toxicity test used. This relationship between the magnitude of the WER and the sensitivity of the toxicity test is due to the aqueous chemistry of metals and is not related to the test organisms or the type of test. The available data indicate that WERs determined with different tests do not differ greatly if the tests have about the same sensitivities, but the data also support the generalization that less sensitive toxicity tests usually give smaller WERs than more sensitive tests (see Appendix D).
- b. When the CCC is lower than the CMC, it is likely that a larger WER will result from tests that are sensitive at the CCC than from tests that are sensitive at the CMC.
- c. The considerations concerning the sensitivities of two tests should also apply to two endpoints for the same test. For any lethality test, use of the LC25 is likely to result in a larger WER than use of the LC50, although the difference might not be measurable in most cases and the LC25 is likely to be more variable than the LC50. Selecting the percent effect to be used to define the endpoint might take into account (a) whether the endpoint is above or below the CMC and/or the CCC and (b) the data obtained when tests are conducted. Once the percent effect is selected for a particular test (e.g., a 48-hr LC50 with 1-day-old fathead minnows), the same percent effect must be used whenever that test is used to determine a WER for that effluent. Similarly, if two different tests with the same species (e.g., a lethality test and a sublethal test) have substantially different sensitivities, both a cmcWER and a cccWER could be obtained with the same species.
- d. The primary toxicity test used in the determination of a WER should have an endpoint in laboratory dilution water that is close to, but not lower than, the CMC and/or CCC to which the WER is to be applied.
- e. Because the endpoint of the primary test in laboratory dilution water cannot be lower than the CMC and/or CCC, the magnitude of the WER is likely to become closer to 1 as the endpoint of the primary test becomes closer to the CMC and/or CCC (see Appendix D).
- f. The WER obtained with the primary test should be confirmed with a secondary test that uses a species that is taxonomically different from the species used in the primary test.
 - 1) The endpoint of the secondary test may be higher or lower than the CMC, the CCC, or the endpoint of the primary test.

- 2) Because of the limited number of toxicity tests that have sensitivities near the CMC or CCC for a metal, it seems unreasonable to require that the two species be further apart taxonomically than being in different orders.

Two different endpoints with the same species **must not** be used as the primary and secondary tests, even if one endpoint is lethal and the other is sublethal.

- g. If more sensitive toxicity tests generally give larger WERs than less sensitive tests, the maximum value of a WER will usually be obtained using a toxicity test whose endpoint in laboratory dilution water equals the CMC or CMC. If such a test is not used, the maximum possible WER probably will not be obtained.
- h. No rationale exists to support the idea that different species or tests with the same sensitivity will produce different WERs. Because the mode of action might differ from species to species and/or from effect to effect, it is easy to speculate that in some cases the magnitude of a WER will depend to some extent on the species, life stage, and/or kind of test, but no data are available to support conclusions concerning the existence and/or magnitude of any such differences.
- i. If the tests are otherwise acceptable, both cmcWERs and cccWERs may be determined using acute and/or chronic tests and using lethal and/or sublethal endpoints. The important consideration is the sensitivity of the test, not the duration, species, life stage, or adverse effect used.
- j. There is no reason to use species that occur at the site; they may be used in the determination of a WER if desired, but:
 - 1) It might be difficult to determine which of the species that occur at the site are sensitive to the metal and are adaptable to laboratory conditions.
 - 2) Species that occur at the site might be harder to obtain in sufficient numbers for conducting toxicity tests over the testing period.
 - 3) Additional QA tests will probably be needed (see section C.3.b) because data are not likely to be available from other laboratories for comparison with the results in laboratory dilution water.
- k. Because a WER is a ratio of results obtained with the same test in two different dilution waters, toxicity tests that are used in WET testing, for example, may be used, even if the national aquatic life guidelines (U.S. EPA 1985) do not allow use of the test in the derivation of an aquatic life criterion. Of course, a test whose endpoint in laboratory dilution water is below the CMC and/or CCC that is to be adjusted cannot be used as a primary test.

- l. Because there is no rationale that suggest that it makes any difference whether the test is conducted with a species that is warmwater or coldwater, a fish or an invertebrate, or resident or nonresident at the site, other than the fact that less sensitive tests are likely to give smaller WERs, such considerations as the availability of test organisms might be important in the selection of the test. Information in Appendix I, a criteria document for the metal of concern (see Appendix E), or any other pertinent source might be useful when selecting primary and secondary tests.
- m. A test in which the test organisms are not fed might give a different WER than a test in which the organisms are fed just because of the presence of the food (see Appendix D). This might depend on the metal, the type and amount of food, and whether a total recoverable or dissolved WER is determined.

Different tests with similar sensitivities are expected to give similar WERs, except for experimental variation. The purpose of the secondary test is to provide information concerning this assumption and the validity of the WER.

4. Should a cmcWER or a cccWER or both be determined?

This question does not have to be answered if the criterion for the site contains either a CMC or a CCC but not both. For example, a body of water that is protected for put-and-take fishing might have only a CMC, whereas a stream whose design flow is zero might have only a CCC.

When the criterion contains both a CMC and a CCC, the simplistic way to answer the question is to determine whether the CMC or the CCC controls the existing permit limits; which one is controlling depends on (a) the ratio of the CMC to the CCC, (b) whether the number of mixing zones is zero, one, or two, and (c) which steady-state or dynamic model was used in the calculation of the permit limits. A better way to answer the question would be to also determine how much the controlling value would have to be changed for the other value to become controlling; this might indicate that it would not be cost-effective to derive, for example, a site-specific CMC (ssCMC) without also deriving a site-specific CCC (ssCCC). There are also other possibilities: (1) It might be appropriate to use a phased approach, i.e., determine either the cmcWER or the cccWER and then decide whether to determine the other. (2) It might be appropriate and environmentally conservative to determine a WER that can be applied to both the CMC and the CCC. (3) It is always allowable to determine and use both a cmcWER and a cccWER, although both can be determined only if toxicity tests with appropriate sensitivities are available.

Because the phased approach can always be used, it is only important to decide whether to use a different approach when its use might be cost-effective. Deciding whether to use a different approach and selecting which one to use is complex because a number of considerations need to be taken into account:

- a. Is the CMC equal to or higher than the CCC?

If the CMC equals the CCC, two WERs cannot be determined if they would be determined using the same site water, but two WERs could be determined if the cmcWER and the cccWER would be determined using different site waters, e.g., waters that contain different concentrations of the effluent.

- b. If the CMC is higher than the CCC, is there a toxicity test whose endpoint in laboratory dilution water is between the CMC and the CCC?

If the CMC is higher than the CCC and there is a toxicity test whose endpoint in laboratory dilution water is between the CMC and the CCC, both a cmcWER and a cccWER can be determined. If the CMC is higher than the CCC but no toxicity test has an endpoint in laboratory dilution water between the CMC and the CCC, two WERs cannot be determined if they would be determined using the same site water; two WERs could be determined if they were determined using different site waters, e.g., waters that contain different concentrations of the effluent.

- c. Was a steady-state or a dynamic model used in the calculation of the permit limits?

It is complex, but reasonably clear, how to make a decision when a steady-state model was used, but it is not clear how a decision should be made when a dynamic model was used.

- d. If a steady-state model was used, were one or two design flows used, i.e., was the hydrologically based steady-state method used or was the biologically based steady-state method used?

When the hydrologically based method is used, one design flow is used for both the CMC and the CCC, whereas when the biologically based method is used, there is a CMC design flow and a CCC design flow. When WERs are determined using downstream water, use of the biologically based method will probably cause the percent effluent in the site water used in the determination of the cmcWER to be different from the percent effluent in the site water used in the determination of the cccWER; thus the two WERs should be determined using two different site waters. This does not impact WERs determined using upstream water.

e. Is there an acute mixing zone? Is there a chronic mixing zone?

1. When WERs are determined using upstream water, the presence or absence of mixing zones has no impact; the cmcWER and the cccWER will both be determined using site water that contains zero percent effluent, i.e., the two WERs will be determined using the same site water.
2. Even when downstream water is used, whether there is an acute mixing zone affects the point of application of the CMC or ssCMC, but it does not affect the determination of any WER.
3. The existence of a chronic mixing zone has important implications for the determination of WERs when downstream water is used (see Appendix A). When WERs are determined using downstream water, the cmcWER should be determined using water at the edge of the chronic mixing zone, whereas the cccWER should be determined using water from a complete-mix situation. (If the biologically based method is used, the two different design flows should also be taken into account when determining the percent effluent that should be in the simulated downstream water.) Thus the percent effluent in the site water used in the determination of the cmcWER will be different from the percent effluent in the site water used in the determination of the cccWER; this is important because the magnitude of a WER will often depend substantially on the percent effluent in the water (see Appendix D).

f. In what situations would it be environmentally conservative to determine one WER and use it to adjust both the cmcWER and the cccWER?

Because (1) the CMC is never lower than the CCC and (2) a more sensitive test will generally give a WER closer to 1, it will be environmentally conservative to use a cmcWER to adjust a CCC when there are no contradicting considerations. In this case, a cmcWER can be determined and used to adjust both the CMC and the CCC. Because water quality can affect the WER, this approach is necessarily valid only if the cmcWER and the cccWER are determined in the same site water. Other situations in which it would be environmentally conservative to use one WER to adjust both the CMC and the CCC are described below.

These considerations have one set of implications when both the cmcWER and cccWER are to be determined using the same site water, and another set of implications when the two WERs are to be determined using different site waters, e.g., when the site waters contain different concentrations of effluent.

When WERs are determined using upstream water, the same site water is used in the determination of both the cmcWER and the cccWER. Whenever the two WERs are determined in the same site water, any difference in the magnitude of the cmcWER and the cccWER will probably be due to the sensitivities of the toxicity tests used. Therefore:

- a. If more sensitive toxicity tests generally give larger WERs than less sensitive tests, the maximum cccWER (a cccWER determined with a test whose endpoint equals the CCC) will usually be larger than the maximum cmcWER because the CCC is never higher than the CMC.
- b. Because the CCC is never higher than the CMC, the maximum cmcWER will usually be smaller than the maximum cccWER and it will be environmentally conservative to use the cmcWER to adjust the CCC.
- c. A cccWER can be determined separately from a cmcWER only if there is a toxicity test with an endpoint in laboratory dilution water that is between the CMC and the CCC. If no such test exists or can be devised, only a cmcWER can be determined, but it can be used to adjust both the CMC and the CCC.
- d. Unless the experimental variation is increased, use of a cccWER, instead of a cmcWER, to adjust the CCC will usually improve the accuracy of the resulting site-specific CCC. Thus a cccWER may be determined and used whenever desired, if a toxicity test has an endpoint in laboratory dilution water between the CMC and the CCC.
- e. A cccWER cannot be used to adjust a CMC if the cccWER was determined using an endpoint that was lower than the CMC in laboratory dilution water because it will probably reduce the level of protection.
- f. Even if there is a toxicity test that has an endpoint in laboratory dilution water that is between the CMC and the CCC, it is not necessary to decide initially whether to determine a cmcWER and/or a cccWER. When upstream water is used, it is always allowable to determine a cmcWER and use it to derive a site-specific CMC and a site-specific CCC and then decide whether to determine a cccWER.
- g. If there is a toxicity test whose endpoint in laboratory dilution water is between the CCC and the CMC, and if this test is used as the secondary test in the determination of the cmcWER, this test will provide information that should be very useful for deciding whether to determine a cccWER in addition to a cmcWER. Further, if it is decided to determine a cccWER, the same two tests used in the determination of the cmcWER could then be used in the determination of the cccWER, with a reversal of their roles as primary and secondary tests. Alternatively, a cmcWER and a cccWER could be determined simultaneously if both tests are conducted on each sample of site water.

When WERs are determined using downstream water, the magnitude of each WER will probably depend on the concentration of effluent in the downstream water used (see Appendix D). The first important consideration is whether the design flow is greater than zero, and the second is whether there is a chronic mixing zone.

- a. If the design flow is zero, cmcWERS and/or cccWERS that are determined for design-flow conditions will both be determined in 100 percent effluent. Thus this case is similar to using upstream water in that both WERs are determined in the same site water. When WERs are determined for high-flow conditions, it will make a difference whether a chronic mixing zone needs to be taken into account, which is the second consideration.
- b. If there is no chronic mixing zone, both WERs will be determined for the complete-mix situation; this case is similar to using upstream water in that both WERs are determined using the same site water. If there is a chronic mixing zone, cmcWERS should be determined in the site water that exists at the edge of the chronic mixing zone, whereas cccWERS should be determined for the complete-mix situation (see Appendix A). Thus the percent effluent will be higher in the site water used in the determination of the cmcWER than in the site water used in the determination of the cccWER. Because a site water with a higher percent effluent will probably give a larger WER than a site water with a lower percent effluent, both a cmcWER and a cccWER can be determined even if there is no test whose endpoint in laboratory dilution water is between the CMC and the CCC. There are opposing considerations, however:
 - 1) The site water used in the determination of the cmcWER will probably have a higher percent effluent than the site water used in the determination of the cccWER, which will tend to cause the cmcWER to be larger than the cccWER.
 - 2) If there is a toxicity test whose endpoint in laboratory dilution water is between the CMC and the CCC, use of a more sensitive test in the determination of the cccWER will tend to cause the cccWER to be larger than the cmcWER.

One consequence of these opposing considerations is that it is not known whether use of the cmcWER to adjust the CCC would be environmentally conservative; if this simplification is not known to be conservative, it should not be used. Thus it is important whether there is a toxicity test whose endpoint in laboratory dilution water is between the CMC and the CCC:

- a. If no toxicity test has an endpoint in laboratory dilution water between the CMC and the CCC, the two WERs have to be determined with the same test, in which case the cmcWER will probably be larger because the

- percent effluent in the site water will be higher. Because of the difference in percent effluent in the site waters that should be used in the determinations of the two WERs, use of the cmcWER to adjust the CCC would not be environmentally conservative, but use of the cccWER to adjust the CMC would be environmentally conservative. Although both WERs could be determined, it would also be acceptable to determine only the cccWER and use it to adjust both the CMC and the CCC.
- b. If there is a toxicity test whose endpoint in laboratory dilution water is between the CMC and the CCC, the two WERs could be determined using different toxicity tests. An environmentally conservative alternative to determining two WERs would be to determine a hybrid WER by using (1) a toxicity test whose endpoint is above the CMC (i.e., a toxicity test that is appropriate for the determination of a cmcWER) and (2) site water for the complete-mix situation (i.e., site water appropriate for the determination of cccWER). It would be environmentally conservative to use this hybrid WER to adjust the CMC and it would be environmentally conservative to use this hybrid WER to adjust the CCC. Although both WERs could be determined, it would also be acceptable to determine only the hybrid WER and use it to adjust both the CMC and the CCC. (This hybrid WER described here in paragraph b is the same as the cccWER described in paragraph a above in which no toxicity test had an endpoint in laboratory dilution water between the CMC and the CCC.)

5. How should a FWER be derived?

Background

Because of experimental variation and variation in the composition of surface waters and effluents, a single determination of a WER does not provide sufficient information to justify adjustment of a criterion. After a sufficient number of WERs have been determined in an acceptable manner, a Final Water-Effect Ratio (FWER) is derived from the WERs, and the FWER is then used to calculate the site-specific criterion. If both a site-specific CMC and a site-specific CCC are to be derived, both a cmcFWER and a cccFWER have to be derived, unless an environmentally conservative estimate is used in place of the cmcFWER and/or the cccFWER.

When a WER is determined using upstream water, the two major sources of variation in the WER are (a) variability in the quality of the upstream water, much of which might be related to season and/or flow, and (b) experimental

variation. When a WER is determined in downstream water, the four major sources of variation are (a) variability in the quality of the upstream water, much of which might be related to season and/or flow, (b) experimental variation, (c) variability in the composition of the effluent, and (d) variability in the percent effluent in the downstream water. Variability and the possibility of mistakes and rare events make it necessary to try to compromise between (1) providing a high probability of adequate protection and (2) placing too much reliance on the smallest experimentally determined WER, which might reflect experimental variation, a mistake, or a rare event rather than a meaningful difference in the WER.

Various ways can be employed to address variability:

- a. Replication can be used to reduce the impact of some sources of variation and to verify the importance of others.
- b. Because variability in the composition of the effluent might contribute substantially to the variability of the WER, it might be desirable to obtain and store two or more samples of the effluent at slightly different times, with the selection of the sampling times depending on such characteristics of the discharge as the average retention time, in case an unusual WER is obtained with the first sample used.
- c. Because of the possibility of mistakes and rare events, samples of effluent and upstream water should be large enough that portions can be stored for later testing or analyses if an unusual WER is obtained.
- d. It might be possible to reduce the impact of the variability in the percent effluent in the downstream water by establishing a relationship between the WER and the percent effluent.

Confounding of the sources can be a problem when more than one source contributes substantial variability.

When permit limits are calculated using a steady-state model, the limits are based on a design flow, e.g., the 7Q10. It is usually assumed that a concentration of metal in an effluent that does not cause unacceptable effects at the design flow will not cause unacceptable effects at higher flows because the metal is diluted by the increased flow of the upstream water. Decreased protection might occur, however, if an increase in flow increases toxicity more than it dilutes the concentration of metal. When permit limits are based on a national criterion, it is often assumed that the criterion is sufficiently conservative that an increase in toxicity will not be great enough to overwhelm the combination of dilution and the assumed conservatism, even though it is likely that the national criterion is not overprotective of all bodies

of water. When WERs are used to reduce the assumed conservatism, there is more concern about the possibility of increased toxicity at flows higher than the design flow and it is important to (1) determine some WERs that correspond to higher flows or (2) provide some conservatism. If the concentration of effluent in the downstream water decreases as flow increases, WERs determined at higher flows are likely to be smaller than WERs determined at design flow but the concentration of metal will also be lower. If the concentration of TSS increases at high flows, however, both the WER and the concentration of metal might increase. If they are determined in an appropriate manner, WERs determined at flows higher than the design flow can be used in two ways:

- a. As environmentally conservative estimates of WERs determined at design flow.
- b. To assess whether WERs determined at design flow will provide adequate protection at higher flows.

In order to appropriately take into account seasonal and flow effects and their interactions, both ways of using high-flow WERs require that the downstream water used in the determination of the WER be similar to that which actually exists during the time of concern. In addition, high-flow WERs can be used in the second way only if the composition of the downstream water is known. To satisfy the requirements that (a) the downstream water used in the determination of a WER be similar to the actual water and (b) the composition of the downstream water be known, it is necessary to obtain samples of effluent and upstream water at the time of concern and to prepare a simulated downstream water by mixing the samples at the ratio of the flows of the effluent and the upstream water that existed when the samples were obtained.

For the first way of using high-flow WERs, they are used directly as environmentally conservative estimates of the design-flow WER. For the second way of using high-flow WERs, each is used to calculate the highest concentration of metal that could be in the effluent without causing the concentration of metal in the downstream water to exceed the site-specific criterion that would be derived for that water using the experimentally determined WER. This highest concentration of metal in the effluent (HCME) can be calculated as:

$$HCME = \frac{[(CCC)(WER)(eFLOW + uFLOW)] - [(uCONC)(uFLOW)]}{eFLOW}$$

where:

CCC = the national, state, or recalculated CCC (or CMC) that is to be adjusted.

eFLOW = the flow of the effluent that was the basis of the preparation of the simulated downstream water. This should be the flow of the effluent that existed when the samples were taken.

uFLOW = the flow of the upstream water that was the basis of the preparation of the simulated downstream water. This should be the flow of the upstream water that existed when the samples were taken.

uCONC = the concentration of metal in the sample of upstream water used in the preparation of the simulated downstream water.

In order to calculate a HCME from an experimentally determined WER, the only information needed besides the flows of the effluent and the upstream water is the concentration of metal in the upstream water, which should be measured anyway in conjunction with the determination of the WER.

When a steady-state model is used to derive permit limits, the limits on the effluent apply at all flows; thus, each HCME can be used to calculate the highest WER (hWER) that could be used to derive a site-specific criterion for the downstream water at design flow so that there would be adequate protection at the flow for which the HCME was determined. The hWER is calculated as:

$$hWER = \frac{(HCME)(eFLOW_{df}) + (uCONC_{df})(uFLOW_{df})}{(CCC)(eFLOW_{df} + uFLOW_{df})}$$

The suffix "df" indicates that the values used for these quantities in the calculation of the hWER are those that exist at design-flow conditions. The additional datum needed in order to calculate the hWER is the concentration of metal in upstream water at design-flow conditions; if this is assumed to be zero, the hWER will be environmentally conservative. If a WER is determined when uFLOW equals the design flow, hWER = WER.

The two ways of using WERs determined at flows higher than design flow can be illustrated using the following examples. These examples were formulated using the concept of additivity of WERs (see Appendix G). A WER determined in downstream water consists of two components, one due to the effluent (the eWER) and one due to the upstream water (the uWER). If the eWER and uWER are strictly additive, when WERs are determined at various upstream flows, the downstream WERs can be calculated from the composition of the downstream water (the % effluent and the % upstream water) and the two WERs (the eWER and the uWER) using the equation:

$$WER = \frac{(\% \text{ effluent}) (eWER) + (\% \text{ upstream water}) (uWER)}{100}$$

In the examples below, it is assumed that:

- A site-specific CCC is being derived.
- The national CCC is 2 ug/L.
- The eWER is 40.
- The eWER and uWER are constant and strictly additive.
- The flow of the effluent (eFLOW) is always 10 cfs.
- The design flow of the upstream water (uFLOWdf) is 40 cfs.

Therefore:

$$HCME = \frac{[(2 \text{ ug/L}) (WER) (10 \text{ cfs} + uFLOW)] - [(uCONC) (uFLOW)]}{10 \text{ ug/L}}$$

$$hWER = \frac{(HCME) (10 \text{ cfs}) + (uCONCdf) (40 \text{ cfs})}{(2 \text{ ug/L}) (10 \text{ cfs} + 40 \text{ cfs})}$$

In the first example, the uWER is assumed to be 5 and so the upstream site-specific CCC (ussCCC) = (CCC) (uWER) = (2 ug/L) (5) = 10 ug/L. uCONC is assumed to be 0.4 ug/L, which means that the assimilative capacity of the upstream water is 9.6 ug/L.

eFLOW (cfs)	uFLOW (cfs)	At Complete Mix			HCME (ug/L)	hWER
		% Eff.	% Ups.	WER		
10	40	20.0	80.0	12.000	118.4	12.00
10	63	13.7	86.3	9.795	140.5	14.21
10	90	10.0	90.0	8.500	166.4	16.80
10	190	5.0	95.0	6.750	262.4	26.40
10	490	2.0	98.0	5.700	550.4	55.20
10	990	1.0	99.0	5.350	1030.4	103.20
10	1990	0.5	99.5	5.175	1990.4	199.20

As the flow of the upstream water increases, the WER decreases to a limiting value equal to uWER. Because the assimilative capacity is greater than zero, the HCMES and hWERS increase due to the increased dilution of the effluent. The increase in hWER at higher flows will not allow any use of the assimilative capacity of the upstream water because the allowed concentration of metal in the effluent is controlled by the lowest hWER, which is the design-flow hWER in this example. Any WER determined at a higher flow can be used as an environmentally conservative estimate of the design-flow WER, and the hWERS show that the WER of 12 provides adequate protection at all flows. When uFLOW equals the design flow of 40 cfs, WER = hWER.

In the second example, uWER is assumed to be 1, which means that $ussCCC = 2 \text{ ug/L}$. uCONC is assumed to be 2 ug/L, so that $uCONC = ussCCC$. The assimilative capacity of the upstream water is 0 ug/L.

eFLOW (cfs)	uFLOW (cfs)	At Complete Mix			HCME (ug/L)	hWER
		% Eff.	% Ups.	WER		
10	40	20.0	80.0	8.800	80.00	8.800
10	63	13.7	86.3	6.343	80.00	8.800
10	90	10.0	90.0	4.900	80.00	8.800
10	190	5.0	95.0	2.950	80.00	8.800
10	490	2.0	98.0	1.780	80.00	8.800
10	990	1.0	99.0	1.390	80.00	8.800
10	1990	0.5	99.5	1.195	80.00	8.800

All the WERs in this example are lower than the comparable WERs in the first example because the uWER dropped from 5 to 1; the limiting value of the WER at very high flow is 1. Also, the HCMEs and hWERs are independent of flow because the increased dilution does not allow any more metal to be discharged when $uCONC = ussCCC$, i.e., when the assimilative capacity is zero. As in the first example, any WER determined at a flow higher than design flow can be used as an environmentally conservative estimate of the design-flow WER and the hWERs show that the WER of 8.8 determined at design flow will provide adequate protection at all flows for which information is available. When uFLOW equals the design flow of 40 cfs, $WER = hWER$.

In the third example, uWER is assumed to be 2, which means that $ussCCC = 4 \text{ ug/L}$. uCONC is assumed to be 1 ug/L; thus the assimilative capacity of the upstream water is 3 ug/L.

eFLOW (cfs)	uFLOW (cfs)	At Complete Mix			HCME (ug/L)	hWER
		% Eff.	% Ups.	WER		
10	40	20.0	80.0	9.600	92.0	9.60
10	63	13.7	86.3	7.206	98.9	10.29
10	90	10.0	90.0	5.800	107.0	11.10
10	190	5.0	95.0	3.900	137.0	14.10
10	490	2.0	98.0	2.760	227.0	23.10
10	990	1.0	99.0	2.380	377.0	38.10
10	1990	0.5	99.5	2.190	677.0	68.10

All the WERs in this example are intermediate between the comparable WERs in the first two examples because the uWER is now 2, which is between 1 and 5; the limiting value of the WER at very high flow is 2. As in the other examples, any WER determined at a flow higher than design flow can be used as an environmentally conservative estimate of the

design-flow WER and the hWERs show that the WER of 9.6 determined at design flow will provide adequate protection at all flows for which information is available. When uFLOW equals the design flow of 40 cfs, WER = hWER.

If this third example is assumed to be subject to acidic snowmelt in the spring so that the eWER and uWER are less-than-additive and result in a WER of 4.8 (rather than 5.8) at a uFLOW of 90 cfs, the third HCME would be 87 ug/L, and the third hWER would be 9.1. This hWER is lower than the design-flow WER of 9.6, so the site-specific criterion would have to be derived using the WER of 9.1, rather than the design-flow WER of 9.6, in order to provide the intended level of protection. If the eWER and uWER were less-than-additive only to the extent that the third WER was 5.3, the third HCME would be 97 ug/L and the third hWER would be 10.1. In this case, dilution by the increased flow would more than compensate for the WERs being less-than-additive, so that the design-flow WER of 9.6 would provide adequate protection at a uFLOW of 90 cfs. Auxiliary information might indicate whether an unusual WER is real or is an accident; for example, if the hardness, alkalinity, and pH of snowmelt are all low, this information would support a low WER.

If the eWER and uWER were more-than-additive so that the third WER was 10, this WER would not be an environmentally conservative estimate of the design-flow WER. If a WER determined at a higher flow is to be used as an estimate of the design-flow WER and there is reason to believe that the eWER and the uWER might be more-than-additive, a test for additivity can be performed (see Appendix G).

Calculating HCMEs and hWERs is straightforward if the WERs are based on the total recoverable measurement. If they are based on the dissolved measurement, it is necessary to take into account the percent of the total recoverable metal in the effluent that becomes dissolved in the downstream water.

To ensure adequate protection, a group of WERs should include one or more WERs corresponding to flows near the design flow, as well as one or more WERs corresponding to higher flows.

- a. Calculation of hWERs from WERs determined at various flows and seasons identifies the highest WER that can be used in the derivation of a site-specific criterion and still provide adequate protection at all flows for which WERs are available. Use of hWERs eliminates the need to assume that WERs determined at design flow will provide adequate protection at higher flows. Because hWERs are calculated to apply at design flow, they

apply to the flow on which the permit limits are based. The lowest of the hWERs ensures adequate protection at all flows, if hWERs are available for a sufficient range of flows, seasons, and other conditions.

- b. Unless additivity is assumed, a WER cannot be extrapolated from one flow to another and therefore it is not possible to predict a design-flow WER from a WER determined at other conditions. The largest WER is likely to occur at design flow because, of the flows during which protection is to be provided, the design flow is the flow at which the highest concentration of effluent will probably occur in the downstream water. This largest WER has to be experimentally determined; it cannot be predicted.

The examples also illustrate that if the concentration of metal in the upstream water is below the site-specific criterion for that water, in the limit of infinite dilution of the effluent with upstream water, there will be adequate protection. The concern, therefore, is for intermediate levels of dilution. Even if the assimilative capacity is zero, as in the second example, there is more concern at the lower or intermediate flows, when the effluent load is still a major portion of the total load, than at higher flows when the effluent load is a minor contribution.

The Options

To ensure adequate protection over a range of flows, two types of WERs need to be determined:

Type 1 WERs are determined by obtaining samples of effluent and upstream water when the downstream flow is between one and two times higher than what it would be under design-flow conditions.

Type 2 WERs are determined by obtaining samples of effluent and upstream water when the downstream flow is between two and ten times higher than what it would be under design-flow conditions.

The only difference between the two types of samples is the downstream flow at the time the samples are taken. For both types of WERs, the samples should be mixed at the ratio of the flows that existed when the samples were taken so that seasonal and flow-related changes in the water quality characteristics of the upstream water are properly related to the flow at which they occurred. The ratio at which the samples are mixed does not have to be the exact ratio that existed when the samples were taken, but the ratio has to be known, which is why simulated downstream water is used. For each Type 1 WER and each Type 2 WER that is determined, a hWER is calculated.

Ideally, sufficient numbers of both types of WERs would be available and each WER would be sufficiently precise and accurate and the Type 1 WERs would be sufficiently similar that the FWER could be the geometric mean of the Type 1 WERs, unless the FWER had to be lowered because of one or more hWERs. If an adequate number of one or both types of WERs is not available, an environmentally conservative WER or hWER should be used as the FWER.

Three Type 1 and/or Type 2 WERs, which were determined using acceptable procedures and for which there were at least three weeks between any two sampling events, must be available in order for a FWER to be derived. If three or more are available, the FWER should be derived from the WERs and hWERs using the lowest numbered option whose requirements are satisfied:

1. If there are two or more Type 1 WERs:
 - a. If at least nineteen percent of all of the WERs are Type 2 WERs, the derivation of the FWER depends on the properties of the Type 1 WERs:
 - 1) If the range of the Type 1 WERs is not greater than a factor of 5 and/or the range of the ratios of the Type 1 WER to the concentration of metal in the simulated downstream water is not greater than a factor of 5, the FWER is the lower of (a) the adjusted geometric mean (see Figure 2) of all of the Type 1 WERs and (b) the lowest hWER.
 - 2) If the range of the Type 1 WERs is greater than a factor of 5 and the range of the ratios of the Type 1 WER to the concentration of metal in the simulated downstream water is greater than a factor of 5, the FWER is the lowest of (a) the lowest Type 1 WER, (b) the lowest hWER, and (c) the geometric mean of all the Type 1 and Type 2 WERs, unless an analysis of the joint probabilities of the occurrences of WERs and metal concentrations indicates that a higher WER would still provide the level of protection intended by the criterion. (EPA intends to provide guidance concerning such an analysis.)
 - b. If less than nineteen percent of all of the WERs are Type 2 WERs, the FWER is the lower of (1) the lowest Type 1 WER and (2) the lowest hWER.
2. If there is one Type 1 WER, the FWER is the lowest of (a) the Type 1 WER, (b) the lowest hWER, and (c) the geometric mean of all of the Type 1 and Type 2 WERs.
3. If there are no Type 1 WERs, the FWER is the lower of (a) the lowest Type 2 WER and (b) the lowest hWER.

If fewer than three WERs are available and a site-specific criterion is to be derived using a WER or a FWER, the WER or FWER has to be assumed to be 1. Examples of deriving FWERS using these options are presented in Figure 3.

The options are designed to ensure that:

- a. The options apply equally well to ordinary flowing waters and to streams whose design flow is zero.
- b. The requirements for deriving the FWER as something other than the lowest WER are not too stringent.
- c. The probability is high that the criterion will be adequately protective at all flows, regardless of the amount of data that are available.
- d. The generation of both types of WERs is encouraged because environmental conservatism is built in if both types of WERs are not available in acceptable numbers.
- e. The amount of conservatism decreases as the quality and quantity of the available data increase.

The requirement that three WERs be available is based on a judgment that fewer WERs will not provide sufficient information. The requirement that at least nineteen percent of all of the available WERs be Type 2 WERs is based on a judgment concerning what constitutes an adequate mix of the two types of WERs: when there are five or more WERs, at least one-fifth should be Type 2 WERs.

Because each of these options for deriving a FWER is expected to provide adequate protection, anyone who desires to determine a FWER can generate three or more appropriate WERs and use the option that corresponds to the WERs that are available. The options that utilize the least useful WERs are expected to provide adequate protection because of the way the FWER is derived from the WERs. It is intended that, on the average, Option 1a will result in the highest FWER, and so it is recommended that data generation should be designed to satisfy the requirements of this option if possible. For example, if two Type 1 WERs have been determined, determining a third Type 1 WER will require use of Option 1b, whereas determining a Type 2 WER will require use of Option 1a.

Calculation of the FWER as an adjusted geometric mean raises three issues:

- a. The level of protection would be greater if the lowest WER, rather than an adjusted mean, were used as the FWER. Although true, the intended level of protection is provided by the national aquatic life criterion derived according to the national guidelines; when sufficient data are available and it is clear how the data should be used, there is no reason to add a substantial margin of safety and thereby change the intended level of protection. Use of an adjusted geometric mean is acceptable if sufficient data are available concerning the WER to demonstrate that the adjusted geometric mean will provide the intended level of protection. Use of the lowest of three or more WERs would be justified, if, for example, the criterion had

been lowered to protect a commercially important species and a WER determined with that species was lower than WERs determined with other species.

- b. The level of protection would be greater if the adjustment was to a probability of 0.95 rather than to a probability of 0.70. As above, the intended level of protection is provided by the national aquatic life criterion derived according to the national guidelines. There is no need to substantially increase the level of protection when site-specific criteria are derived.
- c. It would be easier to use the more common arithmetic mean, especially because the geometric mean usually does not provide much more protection than the arithmetic mean. Although true, use of the geometric mean rather than the arithmetic mean is justified on the basis of statistics and mathematics; use of the geometric mean is also consistent with the intended level of protection. Use of the arithmetic mean is appropriate when the values can range from minus infinity to plus infinity. The geometric mean (GM) is equivalent to using the arithmetic mean of the logarithms of the values. WERs cannot be negative, but the logarithms of WERs can. The distribution of the logarithms of WERs is therefore more likely to be normally distributed than is the distribution of the WERs. Thus, it is better to use the GM of WERs. In addition, when dealing with quotients, use of the GM reduces arguments about the correct way to do some calculations because the same answer is obtained in different ways. For example, if $WER1 = (N1)/(D1)$ and $WER2 = (N2)/(D2)$, then the GM of WER1 and WER2 gives the same value as $[(GM \text{ of } N1 \text{ and } N2)/(GM \text{ of } D1 \text{ and } D2)]$ and also equals the square root of $\{[(N1)(N2)]/[(D1)(D2)]\}$.

Anytime the FWER is derived as the lowest of a series of experimentally determined WERs and/or hWERs, the magnitude of the FWER will depend at least in part on experimental variation. There are at least three ways that the influence of experimental variation on the FWER can be reduced:

- a. A WER determined with a primary test can be replicated and the geometric mean of the replicates used as the value of the WER for that determination. Then the FWER would be the lowest of a number of geometric means rather than the lowest of a number of individual WERs. To be true replicates, the replicate determinations of a WER should not be based on the same test in laboratory dilution water, the same sample of site water, or the same sample of effluent.
- b. If, for example, Option 3 is to be used with three Type 2 WERs and the endpoints of both the primary and

secondary tests in laboratory dilution water are above the CMC and/or CCC to which the WER is to apply, WERs can be determined with both the primary and secondary tests for each of the three sampling times. For each sampling time, the geometric mean of the WER obtained with the primary test and the WER obtained with the secondary test could be calculated; then the lowest of these three geometric means could be used as the FWER. The three WERs cannot consist of some WERs determined with one of the tests and some WERs determined with the other test; similarly the three WERs cannot consist of a combination of individual WERs obtained with the primary and/or secondary tests and geometric means of results of primary and secondary tests.

- c. As mentioned above, because the variability of the effluent might contribute substantially to the variability of the WERs, it might be desirable to obtain and store more than one sample of the effluent when a WER is to be determined in case an unusual WER is obtained with the first sample used.

Examples of the first and second ways of reducing the impact of experimental variation are presented in Figure 4. The availability of these alternatives does not mean that they are necessarily cost-effective.

6. For metals whose criteria are hardness-dependent, at what hardness should WERs be determined?

The issue of hardness bears on such topics as acclimation of test organisms to the site water, adjustment of the hardness of the site water, and how an experimentally determined WER should be used. If all WERs were determined at design-flow conditions, it might seem that all WERs should be determined at the design-flow hardness. Some permit limits, however, are not based on the hardness that is most likely to occur at design flow; in addition, conducting all tests at design-flow conditions provides no information concerning whether adequate protection will be provided at other flows. Thus, unless the hardnesses of the upstream water and the effluent are similar and do not vary with flow, the hardness of the site water will not be the same for all WER determinations.

Because the toxicity tests should be begun within 36 hours after the samples of effluent and upstream water are collected, there is little time to acclimate organisms to a sample-specific hardness. One alternative would be to acclimate the organisms to a preselected hardness and then adjust the hardness of the site water, but adjusting the hardness of the site water might have various effects on the toxicity of the metal due to competitive binding and ionic impacts on the test organisms and on the speciation

of the metal; lowering hardness without also diluting the WER is especially problematic. The least objectionable approach is to acclimate the organisms to a laboratory dilution water with a hardness in the range of 50 to 150 mg/L and then use this water as the laboratory dilution water when the WER is determined. In this way, the test organisms will be acclimated to the laboratory dilution water as specified by ASTM (1993a,b,c,d,e).

Test organisms may be acclimated to the site water for a short time as long as this does not cause the tests to begin more than 36 hours after the samples were collected. Regardless of what acclimation procedure is used, the organisms used for the toxicity test conducted using site water are unlikely to be acclimated as well as would be desirable. This is a general problem with toxicity tests conducted in site water (U.S. EPA 1993a,b,c; ASTM 1993f), and its impact on the results of tests is unknown.

For the practical reasons given above, an experimentally determined WER will usually be a ratio of endpoints determined at two different hardnesses and will thus include contributions from a variety of differences between the two waters, including hardness. The disadvantages of differing hardnesses are that (a) the test organisms probably will not be adequately acclimated to site water and (b) additional calculations will be needed to account for the differing hardnesses; the advantages are that it allows the generation of data concerning the adequacy of protection at various flows of upstream water and it provides a way of overcoming two problems with the hardness equations: (1) it is not known how applicable they are to hardnesses outside the range of 25 to 400 mg/L and (2) it is not known how applicable they are to unusual combinations of hardness, alkalinity, and pH or to unusual ratios of calcium and magnesium.

The additional calculations that are necessary to account for the differing hardnesses will also overcome the shortcomings of the hardness equations. The purpose of determining a WER is to determine how much metal can be in a site water without lowering the intended level of protection. Each experimentally determined WER is inherently referenced to the hardness of the laboratory dilution water that was used in the determination of the WER, but the hardness equation can be used to calculate adjusted WERs that are referenced to other hardnesses for the laboratory dilution water. When used to adjust WERs, a hardness equation for a CMC or CCC can be used to reference a WER to any hardness for a laboratory dilution water, whether it is inside or outside the range of 25 to 400 mg/L, because any inappropriateness in the equation

will be automatically compensated for when the adjusted WER is used in the derivation of a FWER and permit limits.

For example, the hardness equation for the freshwater CMC for copper gives CMCs of 9.2, 18, and 34 ug/L at hardnesses of 50, 100, and 200 mg/L, respectively. If acute toxicity tests with Ceriodaphnia reticulata gave an EC50 of 18 ug/L using a laboratory dilution water with a hardness of 100 mg/L and an EC50 of 532.2 ug/L in a site water, the resulting WER would be 29.57. It can be assumed that, within experimental variation, EC50s of 9.2 and 34 ug/L and WERs of 57.85 and 15.65 would have been obtained if laboratory dilution waters with hardnesses of 50 and 200 mg/L, respectively, had been used, because the EC50 of 532.2 ug/L obtained in the site water does not depend on what water is used for the laboratory dilution water. The WERs of 57.85 and 15.65 can be considered to be adjusted WERs that were extrapolated from the experimentally determined WER using the hardness equation for the copper CMC. If used correctly, the experimentally determined WER and all of the adjusted WERs will result in the same permit limits because they are internally consistent and are all based on the EC50 of 532.2 ug/L that was obtained in site water.

A hardness equation for copper can be used to adjust the WER if the hardness of the laboratory dilution water used in the determination of the WER is in the range of 25 to 400 mg/L (preferably in the range of about 40 to 250 mg/L because most of the data used to derive the equation are in this range). However, the hardness equation can be used to adjust WERs to hardnesses outside the range of 25 to 400 mg/L because the basis of the adjusted WER does not change the fact that the EC50 obtained in site water was 532.2 ug/L. If the hardness of the site water was 16 mg/L, the hardness equation would predict an EC50 of 3.153 ug/L, which would result in an adjusted WER of 168.8. This use of the hardness equation outside the range of 25 to 400 mg/L is valid only if the calculated CMC is used with the corresponding adjusted WER. Similarly, if the hardness of the site water had been 447 mg/L, the hardness equation would predict an EC50 of 72.66 ug/L, with a corresponding adjusted WER of 7.325. If the hardness of 447 mg/L were due to an effluent that contained calcium chloride and the alkalinity and pH of the site water were what would usually occur at a hardness of 50 mg/L rather than 400 mg/L, any inappropriateness in the calculated EC50 of 72.66 ug/L will be compensated for in the adjusted WER of 7.325, because the adjusted WER is based on the EC50 of 532.2 ug/L that was obtained using the site water.

In the above examples it was assumed that at a hardness of 100 mg/L the EC50 for C. reticulata equalled the CMC, which is a very reasonable simplifying assumption. If, however, the WER had been determined with the more resistant Daphnia pulex and EC50s of 50 ug/L and 750 ug/L had been obtained using a laboratory dilution water and a site water, respectively, the CMC given by the hardness equation could not be used as the predicted EC50. A new equation would have to be derived by changing the intercept so that the new equation gives an EC50 of 50 ug/L at a hardness of 100 mg/L; this new equation could then be used to calculate adjusted EC50s, which could then be used to calculate corresponding adjusted WERs:

<u>Hardness</u> <u>(mg/L)</u>	<u>EC50</u> <u>(ug/L)</u>	<u>WER</u>
16	8.894	84.33
50	26.022	28.82
100	50.000*	15.00*
200	96.073	7.81
447	204.970	3.66

The values marked with an asterisk are the assumed experimentally determined values; the others were calculated from these values. At each hardness the product of the EC50 times the WER equals 750 ug/L because all of the WERs are based on the same EC50 obtained using site water. Thus use of the WER allows application of the hardness equation for a metal to conditions to which it otherwise might not be applicable.

HCMEs can then be calculated using either the experimentally determined WER or an adjusted WER as long as the WER is applied to the CMC that corresponds to the hardness on which the WER is based. For example, if the concentration of copper in the upstream water was 1 ug/L and the flows of the effluent and upstream water were 9 and 73 cfs, respectively, when the samples were collected, the HCME calculated from the WER of 15.00 would be:

$$HCME = \frac{(17.73 \text{ ug/L}) (15) (9 + 73 \text{ cfs}) - (1 \text{ ug/L}) (73 \text{ cfs})}{9 \text{ cfs}} = 2415 \text{ ug/L}$$

because the CMC is 17.73 ug/L at a hardness of 100 mg/L. (The value of 17.73 ug/L is used for the CMC instead of 18 ug/L to reduce roundoff error in this example.) If the hardness of the site water was actually 447 ug/L, the HCME could also be calculated using the WER of 3.66 and the CMC of 72.66 ug/L that would be obtained from the CMC hardness equation:

$$HCME = \frac{(72.66 \text{ ug/L}) (3.66) (9 + 73 \text{ cfs}) - (1 \text{ ug/L}) (73 \text{ cfs})}{9 \text{ cfs}} = 2415 \text{ ug/L} .$$

Either WER can be used in the calculation of the HCME as long as the CMC and the WER correspond to the same hardness and therefore to each other, because:

$$(17.73 \text{ ug/L}) (15) = (72.66 \text{ ug/L}) (3.66) .$$

Although the HCME will be correct as long as the hardness, CMC, and WER correspond to each other, the WER used in the derivation of the FWER must be the one that is calculated using a hardness equation to be compatible with the hardness of the site water. If the hardness of the site water was 447 ug/L, the WER used in the derivation of the FWER has to be 3.66; therefore, the simplest approach is to calculate the HCME using the WER of 3.66 and the corresponding CMC of 72.66 ug/L, because these correspond to the hardness of 447 ug/L, which is the hardness of the site water.

In contrast, the hWER should be calculated using the CMC that corresponds to the design hardness. If the design hardness is 50 mg/L, the corresponding CMC is 9.2 ug/L. If the design flows of the effluent and the upstream water are 9 and 20 cfs, respectively, and the concentration of metal in upstream water at design conditions is 1 ug/L, the hWER obtained from the WER determined using the site water with a hardness of 447 mg/L would be:

$$hWER = \frac{(2415 \text{ ug/L}) (9 \text{ cfs}) + (1 \text{ ug/L}) (20 \text{ cfs})}{(9.2 \text{ ug/L}) (9 \text{ cfs} + 20 \text{ cfs})} = 81.54 .$$

None of these calculations provides a way of extrapolating a WER from one site-water hardness to another. The only extrapolations that are possible are from one hardness of laboratory dilution water to another; the adjusted WERs are based on predicted toxicity in laboratory dilution water, but they are all based on measured toxicity in site water. If a WER is to apply to the design flow and the design hardness, one or more toxicity tests have to be conducted using samples of effluent and upstream water obtained under design-flow conditions and mixed at the design-flow ratio to produce the design hardness. A WER that is specifically appropriate to design conditions cannot be based on predicted toxicity in site water; it has to be based on measured toxicity in site water that corresponds to design-flow conditions. The situation is more complicated if the design hardness is not the hardness that is most likely to occur when effluent and upstream water are mixed at the ratio of the design flows.

B. Background Information and Initial Decisions

1. Information should be obtained concerning the effluent and the operating and discharge schedules of the discharger.
2. The spatial extent of the site to which the WER and the site-specific criterion are intended to apply should be defined (see Appendix A). Information concerning tributaries, the plume, and the point of complete mix should be obtained. Dilution models (U.S. EPA 1993d) and dye dispersion studies (Kilpatrick 1992) might provide information that is useful for defining sites for cmcWERS.
3. If the Recalculation Procedure (see Appendix B) is to be used, it should be performed.
4. Pertinent information concerning the calculation of the permit limits should be obtained:
 - a. What are the design flows, i.e., the flow of the upstream water (e.g., 7Q10) and the flow of the effluent that are used in the calculation of the permit limits? (The design flows for the CMC and CCC might be the same or different.)
 - b. Is there a CMC (acute) mixing zone and/or a CCC (chronic) mixing zone?
 - c. What are the dilution(s) at the edge(s) of the mixing zone(s)?
 - d. If the criterion is hardness-dependent, what is the hardness on which the permit limits are based? Is this a hardness that is likely to occur under design-flow conditions?
5. It should be decided whether to determine a cmcWER and/or a cccWER.
6. The water quality criteria document (see Appendix E) that serves as the basis of the aquatic life criterion should be read to identify any chemical or toxicological properties of the metal that are relevant.
7. If the WER is being determined by or for a discharger, it will probably be desirable to decide what is the smallest WER that is desired by the discharger (e.g., the smallest WER that would not require a reduction in the amount of metal discharged). This "smallest desired WER" might be useful when deciding whether to determine a WER. If a WER is determined, this "smallest desired WER" might be useful when selecting the range of concentrations to be tested in the site water.
8. Information should be read concerning health and safety considerations regarding collection and handling of

effluent and surface water samples and conducting toxicity tests (U.S. EPA 1993a; ASTM 1993a). Information should also be read concerning safety and handling of the metallic salt that will be used in the preparation of the stock solution.

9. The proposed work should be discussed with the appropriate regulatory authority (and possibly the Water Management Division of the EPA Regional Office) before deciding how to proceed with the development of a detailed workplan.
10. Plans should be made to perform one or more rangefinding tests in both laboratory dilution water and site water (see section G.7).

C. Selecting Primary and Secondary Tests

1. For each WER (cmcWER and/or cccWER) to be determined, the primary and secondary tests should be selected using the rationale presented in section A.3, the information in Appendix I, the information in the criteria document for the metal (see Appendix E), and any other pertinent information that is available. When a specific test species is not specified, also select the species. Because at least three WERs **must** be determined with the primary test, but only one **must** be determined with the secondary test, selection of the tests might be influenced by the availability of the species (and the life stage in some cases) during the planned testing period.
 - a. The description of a "test" specifies not only the test species and the duration of the test but also the life stage of the species and the adverse effect on which the results are to be based, all of which can have a major impact on the sensitivity of the test.
 - b. The endpoint (e.g., LC50, EC50, IC50) of the primary test in laboratory dilution water should be as close as possible, but it **must not** be below, the CMC and/or CCC to which the WER is to be applied, because for any two tests, the test that has the lower endpoint is likely to give the higher WER (see Appendix D).

NOTE: If both the Recalculation Procedure and a WER are to be used in the derivation of the site-specific criterion, the Recalculation Procedure **must** be completed first because the recalculated CMC and/or CCC **must** be used in the selection of the primary and secondary tests.
 - c. The endpoint (e.g., LC50, EC50, IC50) of the secondary test in laboratory dilution water should be as close as possible, but may be above or below, the CMC and/or CCC to which the WER is to be applied.

- 1) Because few toxicity tests have endpoints close to the CMC and CCC and because the major use of the secondary test is confirmation (see section I.7.b), the endpoint of the secondary test may be below the CMC or CCC. If the endpoint of the secondary test in laboratory dilution water is above the CMC and/or CCC, it might be possible to use the results to reduce the impact of experimental variation (see Figure 4). If the endpoint of the primary test in laboratory dilution water is above the CMC and the endpoint of the secondary test is between the CMC and CCC, it should be possible to determine both a cccWER and a cmcWER using the same two tests.
 - 2) It is often desirable to conduct the secondary test when the first primary test is conducted in case the results are surprising; conducting both tests the first time also makes it possible to interchange the primary and secondary tests, if desired, without increasing the number of tests that need to be conducted. (If results of one or more rangefinding tests are not available, it might be desirable to wait and conduct the secondary test when more information is available concerning the laboratory dilution water and the site water.)
2. The primary and secondary tests must be conducted with species in different taxonomic orders; at least one species must be an animal and, when feasible, one species should be a vertebrate and the other should be an invertebrate. A plant cannot be used if nutrients and/or chelators need to be added to either or both dilution waters in order to determine the WER. It is desirable to use a test and species for which the rate of success is known to be high and for which the test organisms are readily available. (If the WER is to be used with a recalculated CMC and/or CCC, the species used in the primary and secondary tests do not have to be on the list of species that are used to obtain the recalculated CMC and/or CCC.)
3. There are advantages to using tests suggested in Appendix I or other tests of comparable sensitivity for which data are available from one or more other laboratories.
- a. A good indication of the sensitivity of the test is available. This helps ensure that the endpoint in laboratory dilution water is close to the CMC and/or CCC and aids in the selection of concentrations of the metal to be used in the rangefinding and/or definitive toxicity tests in laboratory dilution water. Tests with other species such as species that occur at the site may be used, but it is sometimes more difficult to obtain, hold, and test such species.

- b. When a WER is determined and used, the results of the tests in laboratory dilution water provide the connection between the data used in the derivation of the national criterion and the data obtained in site water, i.e., the results in laboratory dilution water are a vital link in the derivation and use of a WER. It is, therefore, important to be able to judge the quality of the results in laboratory dilution water. Comparison of results with data from other laboratories evaluates all aspects of the test methodology simultaneously, but for the determination of WERs, the most important aspect is the quality of the laboratory dilution water because the dilution water is the most important difference between the two side-by-side tests from which the WER is calculated. Thus, two tests **must** be conducted for which data are available on the metal of concern in a laboratory dilution water from at least one other laboratory. If both the primary and secondary tests are ones for which acceptable data are available from at least one other laboratory, these are the only two tests that have to be conducted. If, however, the primary and/or secondary tests are ones for which no results are already available for the metal of concern from another laboratory, the first or second time a WER is determined at least two additional tests **must** be conducted in the laboratory dilution water in addition to the tests that are conducted for the determination of WERs (see sections F.5 and I.5).
- 1) For the determination of a WER, data are not required for a reference toxicant with either the primary test or the secondary test because the above requirement provides similar data for the metal for which the WER is actually being determined.
 - 2) See Section I.5 concerning interpretation of the results of these tests before additional tests are conducted.

D. Acquiring and Acclimating Test Organisms

1. The test organisms should be obtained, cultured, held, acclimated, fed, and handled as recommended by the U.S. EPA (1993a,b,c) and/or by ASTM (1993a,b,c,d,e). All test organisms **must** be acceptably acclimated to a laboratory dilution water that satisfies the requirements given in sections F.3 and F.4; an appropriate number of the organisms may be randomly or impartially removed from the laboratory dilution water and placed in the site water when it becomes available in order to acclimate the organisms to the site water for a while just before the tests are begun.

2. The organisms used in a pair of side-by-side tests **must** be drawn from the same population and tested under identical conditions.

E. Collecting and Handling Upstream Water and Effluent

1. Upstream water will usually be mixed with effluent to prepare simulated downstream water. Upstream water may also be used as a site water if a WER is to be determined using upstream water in addition to or instead of determining a WER using downstream water. The samples of upstream water **must** be representative; they **must not** be unduly affected by recent runoff events (or other erosion or resuspension events) that cause higher levels of TSS than would normally be present, unless there is particular concern about such conditions.
2. The sample of effluent used in the determination of a WER **must** be representative; it **must** be collected during a period when the discharger is operating normally. Selection of the date and time of sampling of the effluent should take into account the discharge pattern of the discharger. It might be appropriate to collect effluent samples during the middle of the week to allow for reestablishment of steady-state conditions after shutdowns for weekends and holidays; alternatively, if end-of-the-week slug discharges are routine, they should probably be evaluated. As mentioned above, because the variability of the effluent might contribute substantially to the variability of the WERs, it might be desirable to obtain and store more than one sample of the effluent when WERs are to be determined in case an unusual WER is obtained with the first sample used.
3. When samples of site water and effluent are collected for the determination of the WERs with the primary test, there **must** be at least three weeks between one sampling event and the next. It is desirable to obtain samples in at least two different seasons and/or during times of probable differences in the characteristics of the site water and/or effluent.
4. Samples of upstream water and effluent **must** be collected, transported, handled, and stored as recommended by the U.S. EPA (1993a). For example, samples of effluent should usually be composites, but grab samples are acceptable if the residence time of the effluent is sufficiently long. A sufficient volume should be obtained so that some can be stored for additional testing or analyses if an unusual WER is obtained. Samples **must** be stored at 0 to 4°C in the dark with no air space in the sample container.

5. At the time of collection, the flow of both the upstream water and the effluent **must** be either measured or estimated by means of correlation with a nearby U.S.G.S. gauge, the pH of both upstream water and effluent **must** be measured, and samples of both upstream water and effluent should be filtered for measurement of dissolved metals. Hardness, TSS, TOC, and total recoverable and dissolved metal **must** be measured in both the effluent and the upstream water. Any other water quality characteristics, such as total dissolved solids (TDS) and conductivity, that are monitored monthly or more often by the permittee and reported in the Discharge Monitoring Report **must** also be measured. These and the other measurements provide information concerning the representativeness of the samples and the variability of the upstream water and effluent.
6. "Chain of custody" procedures (U.S. EPA 1991b) should be used for all samples of site water and effluent, especially if the data might be involved in a legal proceeding.
7. Tests **must** be begun within 36 hours after the collection of the samples of the effluent and/or the site water, except that tests may be begun more than 36 hours after the collection of the samples if it would require an inordinate amount of resources to transport the samples to the laboratory and begin the tests within 36 hours.
8. If acute and/or chronic tests are to be conducted with daphnids and if the sample of the site water contains predators, the site water **must** be filtered through a 37- μ m sieve or screen to remove predators.

F. Laboratory Dilution Water

1. The laboratory dilution water **must** satisfy the requirements given by U.S. EPA (1993a,b,c) or ASTM (1993a,b,c,d,e). The laboratory dilution water **must** be a ground water, surface water, reconstituted water, diluted mineral water, or dechlorinated tap water that has been demonstrated to be acceptable to aquatic organisms. If a surface water is used for acute or chronic tests with daphnids and if predators are observed in the sample of the water, it **must** be filtered through a 37- μ m sieve or screen to remove the predators. Water prepared by such treatments as deionization and reverse osmosis **must not** be used as the laboratory dilution water unless salts, mineral water, hypersaline brine, or sea salts are added as recommended by U.S. EPA (1993a) or ASTM (1993a).

2. The concentrations of both TOC and TSS **must** be less than 5 mg/L.
3. The hardness of the laboratory dilution water should be between 50 and 150 mg/L and **must** be between 40 and 220 mg/L. If the criterion for the metal is hardness-dependent, the hardness of the laboratory dilution water **must not** be above the hardness of the site water, unless the hardness of the site water is below 50 mg/L.
4. The alkalinity and pH of the laboratory dilution water **must** be appropriate for its hardness; values for alkalinity and pH that are appropriate for some hardnesses are given by U.S. EPA (1993a) and ASTM (1993a); other corresponding values should be determined by interpolation. Alkalinity should be adjusted using sodium bicarbonate, and pH should be adjusted using aeration, sodium hydroxide, and/or sulfuric acid.
5. It would seem reasonable that, before any samples of site water or effluent are collected, the toxicity tests that are to be conducted in the laboratory dilution water for comparison with results of the same tests from other laboratories (see sections C.3.b and I.5) should be conducted. These should be performed at the hardness, alkalinity, and pH specified in sections F.3 and F.4.

G. Conducting Tests

1. There **must** be no differences between the side-by-side tests other than the composition of the dilution water, the concentrations of metal tested, and possibly the water in which the test organisms are acclimated just prior to the beginning of the tests.
2. More than one test using site water may be conducted side-by-side with a test using laboratory dilution water; the one test in laboratory dilution water will be used in the calculation of several WERs, which means that it is very important that that one test be acceptable.
3. Facilities for conducting toxicity tests should be set up and test chambers should be selected and cleaned as recommended by the U.S. EPA (1993a,b,c) and/or ASTM (1993a,b,c,d,e).
4. A stock solution should be prepared using an inorganic salt that is highly soluble in water.
 - a. The salt does not have to be one that was used in tests that were used in the derivation of the national criterion. Nitrate salts are generally acceptable;

- chloride and sulfate salts of many metals are also acceptable (see Appendix J). It is usually desirable to avoid use of a hygroscopic salt. The salt used should meet A.C.S. specifications for reagent-grade, if such specifications are available; use of a better grade is usually not worth the extra cost. No salt should be used until information concerning safety and handling has been read.
- b. The stock solution may be acidified (using metal-free nitric acid) only as necessary to get the metal into solution.
 - c. The same stock solution **must** be used to add metal to all tests conducted at one time.
5. For tests suggested in Appendix I, the appendix presents the recommended duration and whether the static or renewal technique should be used; additional information is available in the references cited in the appendix. Regardless of whether or not or how often test solutions are renewed when these tests are conducted for other purposes, the following guidance applies to all tests that are conducted for the determination of WERs:
- a. The renewal technique **must** be used for tests that last longer than 48 hr.
 - b. If the concentration of dissolved metal decreases by more than 50 % in 48 hours in static or renewal tests, the test solutions **must** be renewed every 24 hours. Similarly, if the concentration of dissolved oxygen becomes too low, the test solutions **must** be renewed every 24 hours. If one test in a pair of tests is a renewal test, both tests **must** be renewal tests.
 - c. When test solutions are to be renewed, the new test solutions **must** be prepared from the original unspiked effluent and water samples that have been stored at 0 to 4°C in the dark with no air space in the sample container.
 - d. The static technique may be used for tests that do not last longer than 48 hours unless the above specifications require use of the renewal technique.
- If a test is used that is not suggested in Appendix I, the duration and technique recommended for a comparable test should be used.
6. Recommendations concerning temperature, loading, feeding, dissolved oxygen, aeration, disturbance, and controls given by the U.S. EPA (1993a,b,c) and/or ASTM (1993a,b,c,d,e) **must** be followed. The procedures that are used **must** be used in both of the side-by-side tests.
7. To aid in the selection of the concentrations of metals that should be used in the test solutions in site water, a static rangefinding test should be conducted for 8 to 96

hours, using a dilution factor of 10 (or 0.1) or 3.2 (or 0.32) increasing from about a factor of 10 below the value of the endpoint given in the criteria document for the metal or in Appendix I of this document for tests with newly hatched fathead minnows. If the test is not in the criteria document and no other data are available, a mean acute value or other data for a taxonomically similar species should be used as the predicted value. This rangefinding test will provide information concerning the concentrations that should be used to bracket the endpoint in the definitive test and will provide information concerning whether the control survival will be acceptable. If dissolved metal is measured in one or more treatments at the beginning and end of the rangefinding test, these data will indicate whether the concentration should be expected to decrease by more than 50 % during the definitive test. The rangefinding test may be conducted in either of two ways:

- a. It may be conducted using the samples of effluent and site water that will be used in the definitive test. In this case, the duration of the rangefinding test should be as long as possible within the limitation that the definitive test must begin within 36 hours after the samples of effluent and/or site water were collected, except as per section E.7.
- b. It may be conducted using one set of samples of effluent and upstream water with the definitive tests being conducted using samples obtained at a later date. In this case the rangefinding test might give better results because it can last longer, but there is the possibility that the quality of the effluent and/or site water might change. Chemical analyses for hardness and pH might indicate whether any major changes occurred from one sample to the next.

Rangefinding tests are especially desirable before the first set of toxicity tests. It might be desirable to conduct rangefinding tests before each individual determination of a WER to obtain additional information concerning the effluent, dilution water, organisms, etc., before each set of side-by-side tests are begun.

8. Several considerations are important in the selection of the dilution factor for definitive tests. Use of concentrations that are close together will reduce the uncertainty in the WER but will require more concentrations to cover a range within which the endpoints might occur. Because of the resources necessary to determine a WER, it is important that endpoints in both dilution waters be obtained whenever a set of side-by-side tests are conducted. Because static and renewal tests can be used to determine WERs, it is relatively easy to use more treatments than would be used in flow-through tests.

The dilution factor for total recoverable metal **must** be between 0.65 and 0.99, and the recommended factor is 0.7. Although factors between 0.75 and 0.99 may be used, their use will probably not be cost-effective. Because there is likely to be more uncertainty in the predicted value of the endpoint in site water, 6 or 7 concentrations are recommended in the laboratory dilution water, and 8 or 9 in the simulated downstream water, at a dilution factor of 0.7. It might be desirable to use even more treatments in the first of the WER determinations, because the design of subsequent tests can be based on the results of the first tests if the site water, laboratory dilution water, and test organisms do not change too much. The cost of adding treatments can be minimized if the concentration of metal is measured only in samples from treatments that will be used in the calculation of the endpoint.

9. Each test **must** contain a dilution-water control. The number of test organisms intended to be exposed to each treatment, including the controls, **must** be at least 20. It is desirable that the organisms be distributed between two or more test chambers per treatment. If test organisms are not randomly assigned to the test chambers, they **must** be assigned impartially (U.S. EPA 1993a; ASTM 1993a) between all test chambers for a pair of side-by-side tests. For example, it is not acceptable to assign 20 organisms to one treatment, and then assign 20 organisms to another treatment, etc. Similarly, it is not acceptable to assign all the organisms to the test using one of the dilution waters and then assign organisms to the test using the other dilution water. The test chambers should be assigned to location in a totally random arrangement or in a randomized block design.
10. For the test using site water, one of the following procedures should be used to prepare the test solutions for the test chambers and the "chemistry controls" (see section H.1):
 - a. Thoroughly mix the sample of the effluent and place the same known volume of the effluent in each test chamber; add the necessary amount of metal, which will be different for each treatment; mix thoroughly; let stand for 2 to 4 hours; add the necessary amount of upstream water to each test chamber; mix thoroughly; let stand for 1 to 3 hours.
 - b. Add the necessary amount of metal to a large sample of the effluent and also maintain an unspiked sample of the effluent; perform serial dilution using a graduated cylinder and the well-mixed spiked and unspiked samples of the effluent; let stand for 2 to 4 hours; add the necessary amount of upstream water to each test chamber; mix thoroughly; let stand for 1 to 3 hours.

- c. Prepare a large volume of simulated downstream water by mixing effluent and upstream water in the desired ratio; place the same known volume of the simulated downstream water in each test chamber; add the necessary amount of metal, which will be different for each treatment; mix thoroughly and let stand for 1 to 3 hours.
 - d. Prepare a large volume of simulated downstream water by mixing effluent and upstream water in the desired ratio; divide it into two portions; prepare a large volume of the highest test concentration of metal using one portion of the simulated downstream water; perform serial dilution using a graduated cylinder and the well-mixed spiked and unspiked samples of the simulated downstream water; let stand for 1 to 3 hours.
- Procedures "a" and "b" allow the metal to equilibrate somewhat with the effluent before the solution is diluted with upstream water.
- 11. For the test using the laboratory dilution water, either of the following procedures may be used to prepare the test solutions for the test chambers and the "chemistry controls" (see section H.1):
 - a. Place the same known volume of the laboratory dilution water in each test chamber; add the necessary amount of metal, which will be different for each treatment; mix thoroughly; let stand for 1 to 3 hours.
 - b. Prepare a large volume of the highest test concentration in the laboratory dilution water; perform serial dilution using a graduated cylinder and the well-mixed spiked and unspiked samples of the laboratory dilution water; let stand for 1 to 3 hours.
 - 12. The test organisms, which have been acclimated as per section D.1, **must** be added to the test chambers for the site-by-side tests at the same time. The time at which the test organisms are placed in the test chambers is defined as the beginning of the tests, which **must** be within 36 hours of the collection of the samples, except as per section E.7.
 - 13. Observe the test organisms and record the effects and symptoms as specified by the U.S. EPA (1993a,b,c) and/or ASTM (1993a,b,c,d,e). Especially note whether the effects, symptoms, and time course of toxicity are the same in the side-by-side tests.
 - 14. Whenever solutions are renewed, sufficient solution should be prepared to allow for chemical analyses.

H. Chemical and Other Measurements

1. To reduce the possibility of contamination of test solutions before or during tests, thermometers and probes for measuring pH and dissolved oxygen **must not** be placed in test chambers that will provide data concerning effects on test organisms or data concerning the concentration of the metal. Thus measurements of pH, dissolved oxygen, and temperature before or during a test **must** be performed either on "chemistry controls" that contain test organisms and are fed the same as the other test chambers or on aliquots that are removed from the test chambers. The other measurements may be performed on the actual test solutions at the beginning and/or end of the test or the renewal.
2. Hardness (in fresh water) or salinity (in salt water), pH, alkalinity, TSS, and TOC **must** be measured on the upstream water, the effluent, the simulated and/or actual downstream water, and the laboratory dilution water. Measurement of conductivity and/or total dissolved solids (TDS) is recommended in fresh water.
3. Dissolved oxygen, pH, and temperature **must** be measured during the test at the times specified by the U.S. EPA (1993a,b,c) and/or ASTM (1993a,b,c,d,e). The measurements **must** be performed on the same schedule for both of the side-by-side tests. Measurements **must** be performed on both the chemistry controls and actual test solutions at the end of the test.
4. Both total recoverable and dissolved metal **must** be measured in the upstream water, the effluent, and appropriate test solutions for each of the tests.
 - a. The analytical measurements should be sufficiently sensitive and precise that variability in analyses will not greatly increase the variability of the WERS. If the detection limit of the analytical method that will be used to determine the metal is greater than one-tenth of the CCC or CMC that is to be adjusted, the analytical method should probably be improved or replaced (see Appendix C). If additional sensitivity is needed, it is often useful to separate the metal from the matrix because this will simultaneously concentrate the metal and remove interferences. Replicate analyses should be performed if necessary to reduce the impact of analytical variability.
 - 1) EPA methods (U.S. EPA 1983b,1991c) should usually be used for both total recoverable and dissolved measurements, but in some cases alternate methods might have to be used in order to achieve the necessary sensitivity. Approval for use of

alternate methods is to be requested from the appropriate regulatory authority.

- b. All measurements of metals **must** be performed using appropriate QA/QC techniques. Clean techniques for obtaining, handling, storing, preparing, and analyzing the samples should be used when necessary to achieve blanks that are sufficiently low (see Appendix C).
- c. Rather than measuring the metal in all test solutions, it is often possible to store samples and then analyze only those that are needed to calculate the results of the toxicity tests. For dichotomous data (e.g., either-or data; data concerning survival), the metal in the following **must** be measured:
 - 1) all concentrations in which some, but not all, of the test organisms were adversely affected.
 - 2) the highest concentration that did not adversely affect any test organisms.
 - 3) the lowest concentration that adversely affected all of the test organisms.
 - 4) the controls.For data that are not dichotomous (i.e., for count and continuous data), the metal in the controls and in the treatments that define the concentration-effect curve **must** be measured; measurement of the concentrations of metals in other treatments is desirable.
- d. In each treatment in which the concentration of metal is to be measured, both the total recoverable and dissolved concentrations **must** be measured:
 - 1) Samples **must** be taken for measurement of total recoverable metal once for a static test, and once for each renewal for renewal tests; in renewal tests, the samples are to be taken after the organisms have been transferred to the new test solutions. When total recoverable metal is measured in a test chamber, the whole solution in the chamber **must** be mixed before the sample is taken for analysis; the solution in the test chamber **must not** be acidified before the sample is taken. The sample **must** be acidified after it is placed in the sample container.
 - 2) Dissolved metal **must** be measured at the beginning and end of each static test; in a renewal test, the dissolved metal **must** be measured at the beginning of the test and just before the solution is renewed the first time. When dissolved metal is measured in a test chamber, the whole solution in the test chamber **must** be mixed before a sufficient amount is removed for filtration; the solution in the test chamber **must not** be acidified before the sample is taken. The sample **must** be filtered within one hour after it is taken, and the filtrate **must** be acidified after filtration.

5. Replicates, matrix spikes, and other QA/QC checks **must** be performed as required by the U.S. EPA (1983a,1991c).

I. Calculating and Interpreting the Results

1. To prevent roundoff error in subsequent calculations, at least four significant digits **must** be retained in all endpoints, WERs, and FWERs. This requirement is not based on mathematics or statistics and does not reflect the precision of the value; its purpose is to minimize concern about the effects of rounding off on a site-specific criterion. All of these numbers are intermediate values in the calculation of permit limits and should not be rounded off as if they were values of ultimate concern.
2. Evaluate the acceptability of each toxicity test individually.
 - a. If the procedures used deviated from those specified above, particularly in terms of acclimation, randomization, temperature control, measurement of metal, and/or disease or disease-treatment, the test should be rejected; if deviations were numerous and/or substantial, the test **must** be rejected.
 - b. Most tests are unacceptable if more than 10 percent of the organisms in the controls were adversely affected, but the limit is higher for some tests; for the tests recommended in Appendix I, the references given should be consulted.
 - c. If an LC50 or EC50 is to be calculated:
 - 1) The percent of the organisms that were adversely affected **must** have been less than 50 percent, and should have been less than 37 percent, in at least one treatment other than the control.
 - 2) In laboratory dilution water the percent of the organisms that were adversely affected **must** have been greater than 50 percent, and should have been greater than 63 percent, in at least one treatment. In site water the percent of the organisms that were adversely affected should have been greater than 63 percent in at least one treatment. (The LC50 or EC50 may be a "greater than" or "less than" value in site water, but not in laboratory dilution water.)
 - 3) If there was an inversion in the data (i.e., if a lower concentration killed or affected a greater percentage of the organisms than a higher concentration), it **must not** have involved more than two concentrations that killed or affected between 20 and 80 percent of the test organisms.

If an endpoint other than an LC50 or EC50 is used or if Abbott's formula is used, the above requirements will have to be modified accordingly.

- d. Determine whether there was anything unusual about the test results that would make them questionable.
 - e. If solutions were not renewed every 24 hours, the concentration of dissolved metal **must not** have decreased by more than 50 percent from the beginning to the end of a static test or from the beginning to the end of a renewal in a renewal test in test concentrations that were used in the calculation of the results of the test.
3. Determine whether the effects, symptoms, and time course of toxicity was the same in the side-by-side tests in the site water and the laboratory dilution water. For example, did mortality occur in one acute test, but immobilization in the other? Did most deaths occur before 24 hours in one test, but after 24 hours in the other? In sublethal tests, was the most sensitive effect the same in both tests? If the effects, symptoms, and/or time course of toxicity were different, it might indicate that the test is questionable or that additivity, synergism, or antagonism occurred in site water. Such information might be particularly useful when comparing tests that produced unusually low or high WERs with tests that produced moderate WERs.
 4. Calculate the results of each test:
 - a. If the data for the most sensitive effect are dichotomous, the endpoint **must** be calculated as a LC50, EC50, LC25, EC25, etc., using methods described by the U.S. EPA (1993a) or ASTM (1993a). If two or more treatments affected between 0 and 100 percent in both tests in a side-by-side pair, probit analysis **must** be used to calculate results of both tests, unless the probit model is rejected by the goodness of fit test in one or both of the acute tests. If probit analysis cannot be used, either because fewer than two percentages are between 0 and 100 percent or because the model does not fit the data, computational interpolation **must** be used (see Figure 5); graphical interpolation **must not** be used.
 - 1) The same endpoint (LC50, EC25, etc.) and the same computational method **must** be used for both tests used in the calculation of a WER.
 - 2) The selection of the percentage used to define the endpoint might be influenced by the percent effect that occurred in the tests and the correspondence with the CCC and/or CMC.
 - 3) If no treatment killed or affected more than 50 percent of the test organisms and the test was otherwise acceptable, the LC50 or EC50 should be reported to be greater than the highest test concentration.

- 4) If no treatment other than the control killed or affected less than 50 percent of the test organisms and the test was otherwise acceptable, the LC50 or EC50 should be reported to be less than the lowest test concentration.
 - b. If the data for the most sensitive effect are not dichotomous, the endpoint **must** be calculated using a regression-type method (Hoekstra and Van Ewijk 1993; Stephan and Rogers 1985), such as linear interpolation (U.S. EPA 1993b,c) or a nonlinear regression method (Barnthouse et al. 1987; Suter et al. 1987; Bruce and Versteeg 1992). The selection of the percentage used to define the endpoint might be influenced by the percent effect that occurred in the tests and the correspondence with the CCC and/or CMC. The endpoints in the side-by-side tests **must** be based on the same amount of the same adverse effect so that the WER is a ratio of identical endpoints. The same computational method **must** be used for both tests used in the calculation of the WER.
 - c. Both total recoverable and dissolved results should be calculated for each test.
 - d. Results should be based on the time-weighted average measured metal concentrations (see Figure 6).
5. The acceptability of the laboratory dilution water **must** be evaluated by comparing results obtained with two sensitive tests using the laboratory dilution water with results that were obtained using a comparable laboratory dilution water in one or more other laboratories (see sections C.3.b and F.5).
- a. If, after taking into account any known effect of hardness on toxicity, the new values for the endpoints of both of the tests are (1) more than a factor of 1.5 higher than the respective means of the values from the other laboratories or (2) more than a factor of 1.5 lower than the respective means of values from the other laboratories or (3) lower than the respective lowest values available from other laboratories or (4) higher than the respective highest values available from other laboratories, the new and old data **must** be carefully evaluated to determine whether the laboratory dilution water used in the WER determination was acceptable. For example, there might have been an error in the chemical measurements, which might mean that the results of all tests performed in the WER determination need to be adjusted and that the WER would not change. It is also possible that the metal is more or less toxic in the laboratory dilution water used in the WER determination. Further, if the new data were based on measured concentrations but the old data were based on nominal concentrations, the new data

should probably be considered to be better than the old. Evaluation of results of any other toxicity tests on the same or a different metal using the same laboratory dilution water might be useful.

- b. If, after taking into account any known effect of hardness on toxicity, the new values for the endpoints of the two tests are not either both higher or both lower in comparison than data from other laboratories (as per section a above) and if both of the new values are within a factor of 2 of the respective means of the previously available values or are within the ranges of the values, the laboratory dilution water used in the WER determination is acceptable.
- c. A control chart approach may be used if sufficient data are available.
- d. If the comparisons do not indicate that the laboratory dilution water, test method, etc., are acceptable, the tests probably should be considered unacceptable, unless other toxicity data are available to indicate that they are acceptable.

Comparison of results of tests between laboratories provides a check on all aspects of the test procedure; the emphasis here is on the quality of the laboratory dilution water because all other aspects of the side-by-side tests on which the WER is based **must** be the same, except possibly for the concentrations of metal used and the acclimation just prior to the beginning of the tests.

- 6. If all the necessary tests and the laboratory dilution water are acceptable, a WER **must** be calculated by dividing the endpoint obtained using site water by the endpoint obtained using laboratory dilution water.
 - a. If both a primary test and a secondary test were conducted using both waters, WERs **must** be calculated for both tests.
 - b. Both total recoverable and dissolved WERs **must** be calculated.
 - c. If the detection limit of the analytical method used to measure the metal is above the endpoint in laboratory dilution water, the detection limit **must** be used as the endpoint, which will result in a lower WER than would be obtained if the actual concentration had been measured. If the detection limit of the analytical method used is above the endpoint in site water, a WER cannot be determined.
- 7. Investigation of the WER.
 - a. The results of the chemical measurements of hardness, alkalinity, pH, TSS, TOC, total recoverable metal, dissolved metal, etc., on the effluent and the upstream water should be examined and compared with previously available values for the effluent and upstream water,

respectively, to determine whether the samples were representative and to get some indication of the variability in the composition, especially as it might affect the toxicity of the metal and the WER, and to see if the WER correlates with one or more of the measurements.

- b. The WERs obtained with the primary and secondary tests should be compared to determine whether the WER obtained with the secondary test confirmed the WER obtained with the primary test. Equally sensitive tests are expected to give WERs that are similar (e.g., within a factor of 3), whereas a test that is less sensitive will probably give a smaller WER than a more sensitive test (see Appendix D). Thus a WER obtained with a primary test is considered confirmed if either or both of the following are true:
 - 1) the WERs obtained with the primary and secondary tests are within a factor of 3.
 - 2) the test, regardless of whether it is the primary or secondary test, that gives a higher endpoint in the laboratory dilution water also gives the larger WER.If the WER obtained with the secondary test does not confirm the WER obtained with the primary test, the results should be investigated. In addition, WERs probably should be determined using both tests the next time samples are obtained and it would be desirable to determine a WER using a third test. It is also important to evaluate what the results imply about the protectiveness of any proposed site-specific criterion.
- c. If the WER is larger than 5, it should be investigated.
 - 1) If the endpoint obtained using the laboratory dilution water was lower than previously reported lowest value or was more than a factor of two lower than an existing Species Mean Acute Value in a criteria document, additional tests in the laboratory dilution water are probably desirable.
 - 2) If a total recoverable WER was larger than 5 but the dissolved WER was not, is the metal one whose WER is likely to be affected by TSS and/or TOC and was the concentration of TSS and/or TOC high? Was there a substantial difference between the total recoverable and dissolved concentrations of the metal in the downstream water?
 - 3) If both the total recoverable and dissolved WERs were larger than 5, is it likely that there is nontoxic dissolved metal in the downstream water?
- d. The adverse effects and the time-course of effects in the side-by-side tests should be compared. If they are different, it might indicate that the site-water test is questionable or that additivity, synergism, or antagonism occurred in the site water. This might be especially important if the WER obtained with the

secondary test did not confirm the WER obtained with the primary test or if the WER was very large or small.

8. If at least one WER determined with the primary test was confirmed by a WER that was simultaneously determined with the secondary test, the cmcFWER and/or the cccFWER should be derived as described in section A.5.
9. All data generated during the determination of the WER should be examined to see if there are any implications for the national or site-specific aquatic life criterion.
 - a. If there are data for a species for which data were not previously available or unusual data for a species for which data were available, the national criterion might need to be revised.
 - b. If the primary test gives an LC50 or EC50 in laboratory dilution water that is the same as the national CMC, the resulting site-specific CMC should be similar to the LC50 that was obtained with the primary test using downstream water. Such relationships might serve as a check on the applicability of the use of WERs.
 - c. If data indicate that the site-specific criterion would not adequately protect a critical species, the site-specific criterion probably should be lowered.

J. Reporting the Results.

A report of the experimental determination of a WER to the appropriate regulatory authority **must** include the following:

1. Name(s) of the investigator(s), name and location of the laboratory, and dates of initiation and termination of the tests.
2. A description of the laboratory dilution water, including source, preparation, and any demonstrations that an aquatic species can survive, grow, and reproduce in it.
3. The name, location, and description of the discharger, a description of the effluent, and the design flows of the effluent and the upstream water.
4. A description of each sampling station, date, and time, with an explanation of why they were selected, and the flows of the upstream water and the effluent at the time the samples were collected.
5. The procedures used to obtain, transport, and store the samples of the upstream water and the effluent.
6. Any pretreatment, such as filtration, of the effluent, site water, and/or laboratory dilution water.
7. Results of all chemical and physical measurements on upstream water, effluent, actual and/or simulated downstream water, and laboratory dilution water, including hardness (or salinity), alkalinity, pH, and concentrations of total recoverable metal, dissolved metal, TSS, and TOC.

8. Description of the experimental design, test chambers, depth and volume of solution in the chambers, loading and lighting, and numbers of organisms and chambers per treatment.
9. Source and grade of the metallic salt, and how the stock solution was prepared, including any acids or bases used.
10. Source of the test organisms, scientific name and how verified, age, life stage, means and ranges of weights and/or lengths, observed diseases, treatments, holding and acclimation procedures, and food.
11. The average and range of the temperature, pH, hardness (or salinity), and the concentration of dissolved oxygen (as % saturation and as mg/L) during acclimation, and the method used to measure them.
12. The following **must** be presented for each toxicity test:
 - a. The average and range of the measured concentrations of dissolved oxygen, as % saturation and as mg/L.
 - b. The average and range of the test temperature and the method used to measure it.
 - c. The schedule for taking samples of test solutions and the methods used to obtain, prepare, and store them.
 - d. A summary table of the total recoverable and dissolved concentrations of the metal in each treatment, including all controls, in which they were measured.
 - e. A summary table of the values of the toxicological variable(s) for each treatment, including all controls, in sufficient detail to allow an independent statistical analysis of the data.
 - f. The endpoint and the method used to calculate it.
 - g. Comparisons with other data obtained by conducting the same test on the same metal using laboratory dilution water in the same and different laboratories; such data may be from a criteria document or from another source.
 - h. Anything unusual about the test, any deviations from the procedures described above, and any other relevant information.
13. All differences, other than the dilution water and the concentrations of metal in the test solutions, between the side-by-side tests using laboratory dilution water and site water.
14. Comparison of results obtained with the primary and secondary tests.
15. The WER and an explanation of its calculation.

A report of the derivation of a FWER **must** include the following:

1. A report of the determination of each WER that was determined for the derivation of the FWER; all WERs determined with secondary tests **must** be reported along with all WERs that were determined with the primary test.

2. The design flow of the upstream water and the effluent and the hardness used in the derivation of the permit limits, if the criterion for the metal is hardness-dependent.
3. A summary table **must** be presented that contains the following for each WER that was derived:
 - a. the value of the WER and the two endpoints from which it was calculated.
 - b. the hWER calculated from the WER.
 - c. the test and species that was used.
 - d. the date the samples of effluent and site water were collected.
 - e. the flows of the effluent and upstream water when the samples were taken.
 - f. the following information concerning the laboratory dilution water, effluent, upstream water, and actual and/or simulated downstream water: hardness (salinity), alkalinity, pH, and concentrations of total recoverable metal, dissolved metal, TSS, and TOC.
4. A detailed explanation of how the FWER was derived from the WERs that are in the summary table.

METHOD 2: DETERMINING cccWERS FOR AREAS AWAY FROM PLUMES

Method 2 might be viewed as a simple process wherein samples of site water are obtained from locations within a large body of fresh or salt water (e.g., an ocean or a large lake, reservoir, or estuary), a WER is determined for each sample, and the FWER is calculated as the geometric mean of some or all of the WERs. In reality, Method 2 is not likely to produce useful results unless substantial resources are devoted to planning and conducting the study. Most sites to which Method 2 is applied will have long retention times, complex mixing patterns, and a number of dischargers. Because metals are persistent, the long retention times mean that the sites are likely to be defined to cover rather large areas; thus such sites will herein be referred to generically as "large sites". Despite the differences between them, all large sites require similar special considerations regarding the determination of WERs. Because Method 2 is based on samples of actual surface water (rather than simulated surface water), no sample should be taken in the vicinity of a plume and the method should be used to determine cccWERS, not cmcWERS. If WERs are to be determined for more than one metal, Appendix F should be read.

Method 2 uses many of the same methodologies as Method 1, such as those for toxicity tests and chemical analyses. Because the sampling plan is crucial to Method 2 and the plan has to be based on site-specific considerations, this description of Method 2 will be more qualitative than the description of Method 1.

Method 2 is based on use of actual surface water samples, but use of simulated surface water might provide information that is useful for some purposes:

1. It might be desirable to compare the WERs for two discharges that contain the same metal. This might be accomplished by selecting an appropriate dilution water and preparing two simulated surface waters, one that contains a known concentration of one effluent and one that contains a known concentration of the other effluent. The relative magnitude of the two WERs is likely to be more useful than the absolute values of the WERs themselves.
2. It might be desirable to determine whether the eWER for a particular effluent is additive with the WER of the site water (see Appendix G). This can be studied by determining WERs for several different known concentrations of the effluent in site water.
3. An event such as a rain might affect the WER because of a change in the water quality, but it might also reduce the WER just by dilution of refractory metal or TSS. A proportional decrease in the WER and in the concentration of the metal (such as by dilution of refractory metal) will not result in underprotection; if, however, dilution decreases the WER

proportionally more than it decreases the concentration of metal in the downstream water, underprotection is likely to occur. This is essentially a determination of whether the WER is additive when the effluent is diluted with rain water (see Appendix G).

4. An event that increases TSS might increase the total recoverable concentration of the metal and the total recoverable WER without having much effect on either the dissolved concentration or the dissolved WER.

In all four cases, the use of simulated surface water is useful because it allows for the determination of WERs using known concentrations of effluent.

An important step in the determination of any WER is to define the area to be included in the site. The major principle that should be applied when defining the area is the same for all sites: The site should be neither too small nor too large. If the area selected is too small, permit limits might be unnecessarily controlled by a criterion for an area outside the site, whereas too large an area might unnecessarily incorporate spatial complexities that are not relevant to the discharge(s) of concern and thereby unnecessarily increase the cost of determining the WER. Applying this principle is likely to be more difficult for large sites than for flowing-water sites.

Because WERs for large sites will usually be determined using actual, rather than simulated, surface water, there are five major considerations regarding experimental design and data analysis:

1. Total recoverable WERs at large sites might vary so much across time, location, and depth that they are not very useful. An assumption should be developed that an appropriately defined WER will be much more similar across time, location, and depth within the site than will a total recoverable WER. If such an assumption cannot be used, it is likely that either the FWER will have to be set equal to the lowest WER and be overprotective for most of the site or separate site-specific criteria will have to be derived for two or more sites.
 - a. One assumption that is likely to be worth testing is that the dissolved WER varies much less across time, location, and depth within a site than the total recoverable WER. If the assumption proves valid, a dissolved WER can be applied to a dissolved national water quality criterion to derive a dissolved site-specific water quality criterion that will apply to the whole site.
 - b. A second assumption that might be worth testing is that the WER correlates with a water quality characteristic such as TSS or TOC across time, location, and depth.
 - c. Another assumption that might be worth testing is that the dissolved and/or total recoverable WER is mostly due to

nontoxic metal rather than to a water quality characteristic that reduces toxicity. If this is true and if there is variability in the WER, the WER will correlate with the concentration of metal in the site water. This is similar to the first assumption, but this one can allow use of both total recoverable and dissolved WERs, whereas the first one only allows use of a dissolved WER.

If WERs are too variable to be useful and no way can be found to deal with the variability, additional sampling will probably be required in order to develop a WER and/or a site-specific water quality criterion that is either (a) spatially and/or temporally dependent or (b) constant and environmentally conservative for nearly all conditions.

2. An experimental design should be developed that tests whether the assumption is of practical value across the range of conditions that occur at different times, locations, and depths within the site. Each design has to be formulated individually to fit the specific site. The design should try to take into account the times, locations, and depths at which the extremes of the physical, chemical, and biological conditions occur within the site, which will require detailed information concerning the site. In addition, the experimental design should balance available resources with the need for adequate sampling.
 - a. Selection of the number and timing of sampling events should take into account seasonal, weekly, and daily considerations. Intensive sampling should occur during the two most extreme seasons, with confirmatory sampling during the other two seasons. Selection of the day and time of sample collection should take into account the discharge schedules of the major industrial and/or municipal discharges. For example, it might be appropriate to collect samples during the middle of the week to allow for reestablishment of steady-state conditions after shutdowns for weekends and holidays; alternatively, end-of-the-week slug discharges are routine in some situations. In coastal sites, the tidal cycle might be important if facilities discharge, for example, over a four-hour period beginning at slack high tide. Because the highest concentration of effluent in the surface water probably occurs at ebb tide, determination of WERs using site water samples obtained at this time might result in inappropriately large WERs that would result in underprotection at other times; samples with unusually large WERs might be especially useful for testing assumptions. The importance of each consideration should be determined on a case-by-case basis.
 - b. Selection of the number and locations of stations to be sampled within a sampling event should consider the site as a whole and take into account sources of water and discharges, mixing patterns, and currents (and tides in coastal areas). If the site has been adequately

characterized, an acceptable design can probably be developed using existing information concerning (1) sources of the metal and other pollutants and (2) the spatial and temporal distribution of concentrations of the metal and water quality factors that might affect the toxicity of the metal. Samples should not be taken within or near mixing zones or plumes of dischargers; dilution models (U.S. EPA 1993) and dye dispersion studies (Kilpatrick 1992) can indicate areas that should definitely be avoided. Maps, current charts, hydrodynamic models, and water quality models used to allocate waste loads and derive permit limits are likely to be helpful when determining when and where to obtain site-water samples. Available information might provide an indication of the acceptability of site water for testing selected species. The larger and more complex the site, the greater the number of sampling locations that will be needed.

- c. In addition to determining the horizontal location of each sampling station, the vertical location (i.e., depth) of the sampling point needs to be selected. Known mixing regimes, the presence of vertical stratification of TSS and/or salinity, concentration of metal, effluent plumes, tolerance of test species, and the need to obtain samples of site water that span the range of site conditions should be considered when selecting the depth at which the sample is to be taken. Some decisions concerning depth cannot be made until information is obtained at the time of sampling; for example, a conductivity meter, salinometer, or transmissometer might be useful for determining where and at what depth to collect samples. Turbidity might correlate with TSS and both might relate to the toxicity of the metal in site water; salinity can indicate whether the test organisms and the site water are compatible.

Because each site is unique, specific guidance cannot be given here concerning either the selection of the appropriate number and locations of sampling stations within a site or the frequency of sampling. All available information concerning the site should be utilized to ensure that the times, locations, and depths of samples span the range of water quality characteristics that might affect the toxicity of the metal:

- a. High and low concentrations of TSS.
- b. High and low concentrations of effluents.
- c. Seasonal effects.
- d. The range of tidal conditions in saltwater situations.

The sampling plan should provide the data needed to allow an evaluation of the usefulness of the assumption(s) that the experimental design is intended to test. Statisticians should play a key role in experimental design and data analysis, but professional judgment that takes into account pertinent biological, chemical, and toxicological considerations is at least as important as rigorous statistical analysis when

interpreting the data and determining the degree to which the data correspond to the assumption(s).

3. The details of each sampling design should be formulated with the aid of people who understand the site and people who have a working knowledge of WERs. Because of the complexity of designing a WER study for large sites, the design team should utilize the combined expertise and experience of individuals from the appropriate EPA Region, states, municipalities, dischargers, environmental groups, and others who can constructively contribute to the design of the study. Building a team of cooperating aquatic toxicologists, aquatic chemists, limnologists, oceanographers, water quality modelers, statisticians, individuals from other key disciplines, as well as regulators and those regulated, who have knowledge of the site and the site-specific procedures, is central to success of the derivation of a WER for a large site. Rather than submitting the workplan to the appropriate regulatory authority (and possibly the Water Management Division of the EPA Regional Office) for comment at the end, they should be members of the team from the beginning.
4. Data from one sampling event should always be analyzed prior to the next sampling event with the goal of improving the sampling design as the study progresses. For example, if the toxicity of the metal in surface water samples is related to the concentration of TSS, a water quality characteristic such as turbidity might be measured at the time of collection of water samples and used in the selection of the concentrations to be used in the WER toxicity tests in site water. At a minimum, the team that interprets the results of one sampling event and plans the next should include an aquatic toxicologist, a metals chemist, a statistician, and a modeler or other user of the data.
5. The final interpretation of the data and the derivation of the FWER(s) should be performed by a team. Sufficient data are likely to be available to allow a quantitative estimate of experimental variation, differences between species, and seasonal differences. It will be necessary to decide whether one site-specific criterion can be applied to the whole area or whether separate site-specific criteria need to be derived for two or more sites. The interpretation of the data might produce two or more alternatives that the appropriate regulatory authority could subject to a cost-benefit analysis.

Other aspects of the determination of a WER for a large site are likely to be the same as described for Method 1. For example:

- a. WERs should be determined using two or more sensitive species; the suggestions given in Appendix I should be considered when selecting the tests and species to be used.

- b. Chemical analyses of site water, laboratory dilution water, and test solutions should follow the requirements for the specific test used and those given in this document.
- c. If tests in many surface water samples are compared to one test in a laboratory dilution water, it is very important that that one test be acceptable. Use of (1) rangefinding tests, (2) additional treatments beyond the standard five concentrations plus controls, and (3) dilutions that are functions of the known concentration-effect relationships obtained with the toxicity test and metal of concern will help ensure that the desired endpoints and WERs can be calculated.
- d. Measurements of the concentrations of both total recoverable and dissolved metal should be targeted to the test concentrations whose data will be used in the calculation of the endpoints.
- e. Samples of site water and/or effluent should be collected, handled, and transported so that the tests can begin as soon as is feasible.
- f. If the large site is a saltwater site, the considerations presented in Appendix H ought to be given attention.

Figure 2: Calculating an Adjusted Geometric Mean

Where n = the number of experimentally determined WERs in a set, the "adjusted geometric mean" of the set is calculated as follows:

- Take the logarithm of each of the WERs. The logarithms can be to any base, but natural logarithms (base e) are preferred for reporting purposes.
- Calculate \bar{x} = the arithmetic mean of the logarithms.
- Calculate s = the sample standard deviation of the logarithms:

$$s = \sqrt{\frac{(\bar{x} - \bar{x})^2}{n - 1}}$$

- Calculate SE = the standard error of the arithmetic mean:
 $SE = s/\sqrt{n}$.
- Calculate $A = \bar{x} - (t_{0.7})(SE)$, where $t_{0.7}$ is the value of Student's t statistic for a one-sided probability of 0.70 with $n - 1$ degrees of freedom. The values of $t_{0.7}$ for some common degrees of freedom (df) are:

<u>df</u>	<u>$t_{0.7}$</u>
1	0.727
2	0.617
3	0.584
4	0.569
5	0.559
6	0.553
7	0.549
8	0.546
9	0.543
10	0.542
11	0.540
12	0.539

The values of $t_{0.7}$ for more degrees of freedom are available, for example, on page T-5 of Natrella (1966).

- Take the antilogarithm of A .

This adjustment of the geometric mean accounts for the fact that the means of fifty percent of the sets of WERs are expected to be higher than the actual mean; using the one-sided value of t for 0.70 reduces the percentage to thirty.

Figure 3: An Example Derivation of a FWER

This example assumes that cccWERS were determined monthly using simulated downstream water that was prepared by mixing upstream water with effluent at the ratio that existed when the samples were obtained. Also, the flow of the effluent is always 10 cfs, and the design flow of the upstream water is 40 cfs. (Therefore, the downstream flow at design-flow conditions is 50 cfs.) The concentration of metal in upstream water at design flow is 0.4 ug/L, and the CCC is 2 ug/L. Each FWER is derived from the WERS and hWERS that are available through that month.

Month	eFLOW (cfs)	uFLOW (cfs)	uCONC (ug/L)	WER	HCME (ug/L)	hWER	FWER
March	10	850	0.8	5.2 ^a	826.4	82.80	1.0 ^b
April	10	289	0.6	6.0 ^c	341.5	34.31	1.0 ^b
May	10	300	0.6	5.8 ^c	341.6	34.32	1.0 ^b
June	10	430	0.6	5.7 ^c	475.8	47.74	5.7 ^d
July	10	120	0.4	7.0 ^c	177.2	17.88	5.7 ^d
Aug.	10	85	0.4	10.5 ^e	196.1	19.77	6.80 ^f
Sept.	10	40	0.4	12.0 ^e	118.4	12.00	10.69 ^g
Oct.	10	45	0.4	11.0 ^e	119.2	12.08	10.88 ^g
Nov.	10	150	0.4	7.5 ^c	234.0	23.56	10.88 ^g
Dec.	10	110	0.4	3.5 ^c	79.6	8.12	8.12 ^h
Jan.	10	180	0.6	6.9 ^c	251.4	25.30	8.12 ^h
Feb.	10	244	0.6	6.1 ^c	295.2	29.68	8.12 ^h

- ^a Neither Type 1 nor Type 2; the downstream flow (i.e., the sum of the eFLOW and the uFLOW) is > 500 cfs.
- ^b The total number of available Type 1 and Type 2 WERS is less than 3.
- ^c A Type 2 WER; the downstream flow is between 100 and 500 cfs.
- ^d No Type 1 WER is available; the FWER is the lower of the lowest Type 2 WER and the lowest hWER.
- ^e A Type 1 WER; the downstream flow is between 50 and 100 cfs.
- ^f One Type 1 WER is available; the FWER is the geometric mean of all Type 1 and Type 2 WERS.
- ^g Two or more Type 1 WERS are available and the range is less than a factor of 5; the FWER is the adjusted geometric mean (see Figure 2) of the Type 1 WERS, because all the hWERS are higher.
- ^h Two or more Type 1 WERS are available and the range is not greater than a factor of 5; the FWER is the lowest hWER because the lowest hWER is lower than the adjusted geometric mean of the Type 1 WERS.

Figure 4: Reducing the Impact of Experimental Variation

When the FWER is the lowest of, for example, three WERs, the impact of experimental variation can be reduced by conducting additional primary tests. If the endpoint of the secondary test is above the CMC or CCC to which the FWER is to be applied, the additional tests can also be conducted with the secondary test.

Month	Case 1			Case 2	
	(Primary Test)	(Primary Test)	(Primary Test)	Geometric Mean	
April	4.801	4.801	3.565	4.137	
May	2.552	2.552	4.190	3.270	
June	9.164	9.164	6.736	7.857	
Lowest	2.552			3.270	

Month	Case 3			Case 4		
	(Primary Test)	(Second. Test)	Geo. Mean	(Primary Test)	(Second. Test)	Geo. Mean
April	4.801	3.163	3.897	4.801	3.163	3.897
May	2.552	5.039	3.586	2.552	2.944	2.741
June	9.164	7.110	8.072	9.164	7.110	8.072
Lowest			3.586			2.741

Case 1 uses the individual WERs obtained with the primary test for the three months, and the FWER is the lowest of the three WERs. In Case 2, duplicate primary tests were conducted in each month, so that a geometric mean could be calculated for each month; the FWER is the lowest of the three geometric means.

In Cases 3 and 4, both a primary test and a secondary test were conducted each month and the endpoints for both tests in laboratory dilution water are above the CMC or CCC to which the FWER is to be applied. In both of these cases, therefore, the FWER is the lowest of the three geometric means.

The availability of these alternatives does not mean that they are necessarily cost-effective.

Figure 5: Calculating an LC50 (or EC50) by Interpolation

When fewer than two treatments kill some but not all of the exposed test organisms, a statistically sound estimate of an LC50 cannot be calculated. Some programs and methods produce LC50s when there are fewer than two "partial kills", but such results are obtained using interpolation, not statistics. If (a) a test is otherwise acceptable, (b) a sufficient number of organisms are exposed to each treatment, and (c) the concentrations are sufficiently close together, a test with zero or one partial kill can provide all the information that is needed concerning the LC50. An LC50 calculated by interpolation should probably be called an "approximate LC50" to acknowledge the lack of a statistical basis for its calculation, but this does not imply that such an LC50 provides no useful toxicological information. If desired, the binomial test can be used to calculate a statistically sound probability that the true LC50 lies between two tested concentrations (Stephan 1977).

Although more complex interpolation methods can be used, they will not produce a more useful LC50 than the method described here. Inversions in the data between two test concentrations should be removed by pooling the mortality data for those two concentrations and calculating a percent mortality that is then assigned to both concentrations. Logarithms to a base other than 10 can be used if desired. If P_1 and P_2 are the percentages of the test organisms that died when exposed to concentrations C_1 and C_2 , respectively, and if $C_1 < C_2$, $P_1 < P_2$, $0 \leq P_1 \leq 50$, and $50 \leq P_2 \leq 100$, then:

$$P = \frac{50 - P_1}{P_2 - P_1}$$

$$C = \text{Log } C_1 + P(\text{Log } C_2 - \text{Log } C_1)$$

$$\text{LC50} = 10^C$$

If $P_1 = 0$ and $P_2 = 100$, $\text{LC50} = \sqrt{(C_1)(C_2)}$.

If $P_1 = P_2 = 50$, $\text{LC50} = \sqrt{(C_1)(C_2)}$.

If $P_1 = 50$, $\text{LC50} = C_1$.

If $P_2 = 50$, $\text{LC50} = C_2$.

If $C_1 = 4 \text{ mg/L}$, $C_2 = 7 \text{ mg/L}$, $P_1 = 15 \%$, and $P_2 = 100 \%$, then $\text{LC50} = 5.036565 \text{ mg/L}$.

Besides the mathematical requirements given above, the following toxicological recommendations are given in sections G.8 and I.2:

- $0.65 < C_1/C_2 < 0.99$.
- $0 \leq P_1 < 37$.
- $63 < P_2 \leq 100$.

Figure 6: Calculating a Time-Weighted Average

If a sampling plan (e.g., for measuring metal in a treatment in a toxicity test) is designed so that a series of values are obtained over time in such a way that each value contains the same amount of information (i.e., represents the same amount of time), then the most meaningful average is the arithmetic average. In most cases, however, when a series of values is obtained over time, some values contain more information than others; in these cases the most meaningful average is a time-weighted average (TWA). If each value contains the same amount of information, the arithmetic average will equal the TWA.

A TWA is obtained by multiplying each value by a weight and then dividing the sum of the products by the sum of the weights. The simplest approach is to let each weight be the duration of time that the sample represents. Except for the first and last samples, the period of time represented by a sample starts halfway to the previous sample and ends halfway to the next sample. The period of time represented by the first sample starts at the beginning of the test, and the period of time represented by the last sample ends at the end of the test. Thus for a 96-hr toxicity test, the sum of the weights will be 96 hr.

The following are hypothetical examples of grab samples taken from 96-hr flow-through tests for two common sampling regimes:

<u>Sampling time (hr)</u>	<u>Conc. (mg/L)</u>	<u>Weight (hr)</u>	<u>Product (hr) (mg/L)</u>	<u>Time-weighted average (mg/L)</u>
0	12	48	576	
96	14	<u>48</u>	<u>672</u>	
		96	1248	1248/96 = 13.00
0	8	12	96	
24	6	24	144	
48	7	24	168	
72	9	24	216	
96	8	<u>12</u>	<u>96</u>	
		96	720	720/96 = 7.500

When all the weights are the same, the arithmetic average equals the TWA. Similarly, if only one sample is taken, both the arithmetic average and the TWA equal the value of that sample.

The rules are more complex for composite samples and for samples from renewal tests. In all cases, however, the sampling plan can be designed so that the TWA equals the arithmetic average.

REFERENCES

- ASTM. 1993a. Guide for Conducting Acute Toxicity Tests with Fishes, Macroinvertebrates, and Amphibians. Standard E729. American Society for Testing and Materials, Philadelphia, PA.
- ASTM. 1993b. Guide for Conducting Static Acute Toxicity Tests Starting with Embryos of Four Species of Saltwater Bivalve Molluscs. Standard E724. American Society for Testing and Materials, Philadelphia, PA.
- ASTM. 1993c. Guide for Conducting Renewal Life-Cycle Toxicity Tests with Daphnia magna. Standard E1193. American Society for Testing and Materials, Philadelphia, PA.
- ASTM. 1993d. Guide for Conducting Early Life-Stage Toxicity Tests with Fishes. Standard E1241. American Society for Testing and Materials, Philadelphia, PA.
- ASTM. 1993e. Guide for Conducting Three-Brood, Renewal Toxicity Tests with Ceriodaphnia dubia. Standard E1295. American Society for Testing and Materials, Philadelphia, PA.
- ASTM. 1993f. Guide for Conducting Acute Toxicity Tests on Aqueous Effluents with Fishes, Macroinvertebrates, and Amphibians. Standard E1192. American Society for Testing and Materials, Philadelphia, PA.
- Barnthouse, L.W., G.W. Suter, A.E. Rosen, and J.J. Beauchamp. 1987. Estimating Responses of Fish Populations to Toxic Contaminants. Environ. Toxicol. Chem. 6:811-824.
- Bruce, R.D., and D.J. Versteeg. 1992. A Statistical Procedure for Modeling Continuous Toxicity Data. Environ. Toxicol. Chem. 11:1485-1494.
- Hoekstra, J.A., and P.H. Van Ewijk. 1993. Alternatives for the No-Observed-Effect Level. Environ. Toxicol. Chem. 12:187-194.
- Kilpatrick, F.A. 1992. Simulation of Soluble Waste Transport and Buildup in Surface Waters Using Tracers. Open-File Report 92-457. U.S. Geological Survey, Books and Open-File Reports, Box 25425, Federal Center, Denver, CO 80225.
- Natrella, M.G. 1966. Experimental Statistics. National Bureau of Standards Handbook 91. (Issued August 1, 1963; reprinted October 1966 with corrections). U.S. Government Printing Office, Washington, DC.

Prothro, M.G. 1993. Memorandum titled "Office of Water Policy and Technical Guidance on Interpretation and Implementation of Aquatic Life Metals Criteria". October 1.

Stephan, C.E. 1977. Methods for Calculating an LC50. In: Aquatic Toxicology and Hazard Evaluation. (F.L. Mayer and J.L. Hamelink, eds.) ASTM STP 634. American Society for Testing and Materials, Philadelphia, PA. pp. 65-84.

Stephan, C.E., and J.W. Rogers. 1985. Advantages of Using Regression Analysis to Calculate Results of Chronic Toxicity Tests. In: Aquatic Toxicology and Hazard Assessment: Eighth Symposium. (R.C. Bahner and D.J. Hansen, eds.) ASTM STP 891. American Society for Testing and Materials, Philadelphia, PA. pp. 328-338.

Suter, G.W., A.E. Rosen, E. Linder, and D.F. Parkhurst. 1987. Endpoints for Responses of Fish to Chronic Toxic Exposures. Environ. Toxicol. Chem. 6:793-809.

U.S. EPA. 1983a. Water Quality Standards Handbook. Office of Water Regulations and Standards, Washington, DC.

U.S. EPA. 1983b. Methods for Chemical Analysis of Water and Wastes. EPA-600/4-79-020. National Technical Information Service, Springfield, VA.

U.S. EPA. 1984. Guidelines for Deriving Numerical Aquatic Site-Specific Water Quality Criteria by Modifying National Criteria. EPA-600/3-84-099 or PB85-121101. National Technical Information Service, Springfield, VA.

U.S. EPA. 1985. Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses. PB85-227049. National Technical Information Service, Springfield, VA.

U.S. EPA. 1991a. Technical Support Document for Water Quality-based Toxics Control. EPA/505/2-90-001 or PB91-127415. National Technical Information Service, Springfield, VA.

U.S. EPA. 1991b. Manual for the Evaluation of Laboratories Performing Aquatic Toxicity Tests. EPA/600/4-90/031. National Technical Information Service, Springfield, VA.

U.S. EPA. 1991c. Methods for the Determination of Metals in Environmental Samples. EPA-600/4-91-010. National Technical Information Service, Springfield, VA.

U.S. EPA. 1992. Interim Guidance on Interpretation and Implementation of Aquatic Life Criteria for Metals. Office of Science and Technology, Health and Ecological Criteria Division, Washington, DC.

U.S. EPA. 1993a. Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms. Fourth Edition. EPA/600/4-90/027F. National Technical Information Service, Springfield, VA.

U.S. EPA. 1993b. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms. Third Edition. EPA/600/4-91/002. National Technical Information Service, Springfield, VA.

U.S. EPA. 1993c. Short-Term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms. Second Edition. EPA/600/4-91/003. National Technical Information Service, Springfield, VA.

U.S. EPA. 1993d. Dilution Models for Effluent Discharges. Second Edition. EPA/600/R-93/139. National Technical Information Service, Springfield, VA.

Appendix A: Comparison of WERs Determined Using Upstream and Downstream Water

The "Interim Guidance" concerning metals (U.S. EPA 1992) made a fundamental change in the way WERs should be experimentally determined because it changed the source of the site water. The earlier guidance (U.S. EPA 1983, 1984) required that upstream water be used as the site water, whereas the newer guidance (U.S. EPA 1992) recommended that downstream water be used as the site water. The change in the source of the site water was merely an acknowledgement that the WER that applies at a location in a body of water should, when possible, be determined using the water that occurs at that location.

Because the change in the source of the dilution water was expected to result in an increase in the magnitude of many WERs, interest in and concern about the determination and use of WERs increased. When upstream water was the required site water, it was expected that WERs would generally be low and that the determination and use of WERs could be fairly simple. After downstream water became the recommended site water, the determination and use of WERs was examined much more closely. It was then realized that the determination and use of upstream WERs was more complex than originally thought. It was also realized that the use of downstream water greatly increased the complexity and was likely to increase both the magnitude and the variability of many WERs. Concern about the fate of discharged metal also increased because use of downstream water might allow the discharge of large amounts of metal that has reduced or no toxicity at the end of the pipe. The probable increases in the complexity, magnitude, and variability of WERs and the increased concern about fate, increased the importance of understanding the relevant issues as they apply to WERs determined using both upstream water and downstream water.

A. Characteristics of the Site Water

The idealized concept of an upstream water is a pristine water that is relatively unaffected by people. In the real world, however, many upstream waters contain naturally occurring ligands, one or more effluents, and materials from nonpoint sources; all of these might impact a WER. If the upstream water receives an effluent containing TOC and/or TSS that contributes to the WER, the WER will probably change whenever the quality or quantity of the TOC and/or TSS changes. In such a case, the determination and use of the WER in upstream water will have some of the increased complexity associated with use of downstream water and some of the concerns associated with multiple-discharge situations (see Appendix F). The amount of complexity will depend greatly on the

number and type of upstream point and nonpoint sources, the frequency and magnitude of fluctuations, and whether the WER is being determined above or below the point of complete mix of the upstream sources.

Downstream water is a mixture of effluent and upstream water, each of which can contribute to the WER, and so there are two components to a WER determined in downstream water: the effluent component and the upstream component. The existence of these two components has the following implications:

1. WERs determined using downstream water are likely to be larger and more variable than WERs determined using upstream water.
2. The effluent component should be applied only where the effluent occurs, which has implications concerning implementation.
3. The magnitude of the effluent component of a WER will depend on the concentration of effluent in the downstream water. (A consequence of this is that the effluent component will be zero where the concentration of effluent is zero, which is the point of item 2 above.)
4. The magnitude of the effluent component of a WER is likely to vary as the composition of the effluent varies.
5. Compared to upstream water, many effluents contain higher concentrations of a wider variety of substances that can impact the toxicity of metals in a wider variety of ways, and so the effluent component of a WER can be due to a variety of chemical effects in addition to such factors as hardness, alkalinity, pH, and humic acid.
6. Because the effluent component might be due, in whole or in part, to the discharge of refractory metal (see Appendix D), the WER cannot be thought of simply as being caused by the effect of water quality on the toxicity of the metal.

Dealing with downstream WERs is so much simpler if the effluent WER (eWER) and the upstream WER (uWER) are additive that it is desirable to understand the concept of additivity of WERs, its experimental determination, and its use (see Appendix G).

B. The Implications of Mixing Zones.

When WERs are determined using upstream water, the presence or absence of mixing zones has no impact; the cmcWER and the cccWER will both be determined using site water that contains zero percent of the effluent of concern, i.e., the two WERs will be determined using the same site water.

When WERs are determined using downstream water, the magnitude of each WER will probably depend on the concentration of effluent in the downstream water used (see Appendix D). The concentration of effluent in the site water will depend on

where the sample is taken, which will not be the same for the cmcWER and the cccWER if there are mixing zone(s). Most, if not all, discharges have a chronic (CCC) mixing zone; many, but not all, also have an acute (CMC) mixing zone. The CMC applies at all points except those inside a CMC mixing zone; thus if there is no CMC mixing zone, the CMC applies at the end of the pipe. The CCC applies at all points outside the CCC mixing zone. It is generally assumed that if permit limits are based on a point in a stream at which both the CMC and the CCC apply, the CCC will control the permit limits, although the CMC might control if different averaging periods are appropriately taken into account. For this discussion, it will be assumed that the same design flow (e.g., 7Q10) is used for both the CMC and the CCC.

If the cmcWER is to be appropriate for use inside the chronic mixing zone, but the cccWER is to be appropriate for use outside the chronic mixing zone, the concentration of effluent that is appropriate for use in the determination of the two WERs will not be the same. Thus even if the same toxicity test is used in the determination of the cmcWER and the cccWER, the two WERs will probably be different because the concentration of effluent will be different in the two site waters in which the WERs are determined.

If the CMC is only of concern within the CCC mixing zone, the highest relevant concentration of metal will occur at the edge of the CMC mixing zone if there is a CMC mixing zone; the highest concentration will occur at the end of the pipe if there is no CMC mixing zone. In contrast, within the CCC mixing zone, the lowest cmcWER will probably occur at the outer edge of the CCC mixing zone. Thus the greatest level of protection would be provided if the cmcWER is determined using water at the outer edge of the CCC mixing zone, and then the calculated site-specific CMC is applied at the edge of the CMC mixing zone or at the end of the pipe, depending on whether there is an acute mixing zone. The cmcWER is likely to be lowest at the outer edge of the CCC mixing zone because of dilution of the effluent, but this dilution will also dilute the metal. If the cmcWER is determined at the outer edge of the CCC mixing zone but the resulting site-specific CMC is applied at the end of the pipe or at the edge of the CMC mixing zone, dilution is allowed to reduce the WER but it is not allowed to reduce the concentration of the metal. This approach is environmentally conservative, but it is probably necessary given current implementation procedures. (The situation might be more complicated if the uWER is higher than the eWER or if the two WERs are less-than-additive.)

A comparable situation applies to the CCC. Outside the CCC mixing zone, the CMC and the CCC both apply, but it is assumed that the CMC can be ignored because the CCC will be more

restrictive. The cccWER should probably be determined for the complete-mix situation, but the site-specific CCC will have to be met at the edge of the CCC mixing zone. Thus dilution of the WER from the edge of the CCC mixing zone to the point of complete mix is taken into account, but dilution of the metal is not.

If there is neither an acute nor a chronic mixing zone, both the CMC and the CCC apply at the end of the pipe, but the CCC should still be determined for the complete-mix situation.

C. Definition of site.

In the general context of site-specific criteria, a "site" may be a state, region, watershed, waterbody, segment of a waterbody, category of water (e.g., ephemeral streams), etc., but the site-specific criterion is to be derived to provide adequate protection for the entire site, however the site is defined. Thus, when a site-specific criterion is derived using the Recalculation Procedure, all species that "occur at the site" need to be taken into account when deciding what species, if any, are to be deleted from the dataset. Similarly, when a site-specific criterion is derived using a WER, the WER is to be adequately protective of the entire site. If, for example, a site-specific criterion is being derived for an estuary, WERs could be determined using samples of the surface water obtained from various sampling stations, which, to avoid confusion, should not be called "sites". If all the WERs were sufficiently similar, one site-specific criterion could be derived to apply to the whole estuary. If the WERs were sufficiently different, either the lowest WER could be used to derive a site-specific criterion for the whole estuary, or the data might indicate that the estuary should be divided into two or more sites, each with its own criterion.

The major principle that should be applied when defining the area to be included in the site is very simplistic: The site should be neither too small nor too large.

1. Small sites are probably appropriate for cmcWERS, but usually are not appropriate for cccWERS because metals are persistent, although some oxidation states are not persistent and some metals are not persistent in the water column. For cccWERS, the smaller the defined site, the more likely it is that the permit limits will be controlled by a criterion for an area that is outside the site, but which could have been included in the site without substantially changing the WER or increasing the cost of determining the WER.
2. Too large an area might unnecessarily increase the cost of determining the WER. As the size of the site increases,

- the spatial and temporal variability is likely to increase, which will probably increase the number of water samples in which WERs will need to be determined before a site-specific criterion can be derived.
3. Events that import or resuspend TSS and/or TOC are likely to increase the total recoverable concentration of the metal and the total recoverable WER while having a much smaller effect on the dissolved concentration and the dissolved WER. Where the concentration of dissolved metal is substantially more constant than the concentration of total recoverable metal, the site can probably be much larger for a dissolved criterion than for a total recoverable criterion. If one criterion is not feasible for the whole area, it might be possible to divide it into two or more sites with separate total recoverable or dissolved criteria or to make the criterion dependent on a water quality characteristic such as TSS or salinity.
 4. Unless the site ends where one body of water meets another, at the outer edge of the site there will usually be an instantaneous decrease in the allowed concentration of the metal in the water column due to the change from one criterion to another, but there will not be an instantaneous decrease in the actual concentration of metal in the water column. The site has to be large enough to include the transition zone in which the actual concentration decreases so that the criterion outside the site is not exceeded.

It is, of course, possible in some situations that relevant distant conditions (e.g., a lower downstream pH) will necessitate a low criterion that will control the permit limits such that it is pointless to determine a WER.

When a WER is determined in upstream water, it is generally assumed that a downstream effluent will not decrease the WER. It is therefore assumed that the site can usually cover a rather large geographic area.

When a site-specific criterion is derived based on WERs determined using downstream water, the site should not be defined in the same way that it would be defined if the WER were determined using upstream water. The eWER should be allowed to affect the site-specific criterion wherever the effluent occurs, but it should not be allowed to affect the criterion in places where the effluent does not occur. In addition, insofar as the magnitude of the effluent component at a point in the site depends on the concentration of effluent, the magnitude of the WER at a particular point will depend on the concentration of effluent at that point. To the extent that the eWER and the uWER are additive, the WER and the concentration of metal in the plume will decrease proportionally (see Appendix G).

When WERs are determined using downstream water, the following considerations should be taken into account when the site is defined:

1. If a site-specific criterion is derived using a WER that applies to the complete-mix situation, the upstream edge of the site to which this criterion applies should be the point at which complete mix actually occurs. If the site to which the complete-mix WER is applied starts at the end of the pipe and extends all the way across the stream, there will be an area beside the plume that will not be adequately protected by the site-specific criterion.
2. Upstream of the point of complete mix, it will usually be protective to apply a site-specific criterion that was derived using a WER that was determined using upstream water.
3. The plume might be an area in which the concentration of metal could exceed a site-specific criterion without causing toxicity because of simultaneous dilution of the metal and the eWER. The fact that the plume is much larger than the mixing zone might not be important if there is no toxicity within the plume. As long as the concentration of metal in 100 % effluent does not exceed that allowed by the additive portion of the eWER, from a toxicological standpoint neither the size nor the definition of the plume needs to be of concern because the metal will not cause toxicity within the plume. If there is no toxicity within the plume, the area in the plume might be like a traditional mixing zone in that the concentration of metal exceeds the site-specific criterion, but it would be different from a traditional mixing zone in that the level of protection is not reduced.

Special considerations are likely to be necessary in order to take into account the eWER when defining a site related to multiple discharges. (see Appendix F).

D. The variability in the experimental determination of a WER.

When a WER is determined using upstream water, the two major sources of variation in the WER are (a) variability in the quality of the site water, which might be related to season and/or flow, and (b) experimental variation. Ordinary day-to-day variation will account for some of the variability, but seasonal variation is likely to be more important.

As explained in Appendix D, variability in the concentration of nontoxic dissolved metal will contribute to the variability of both total recoverable WERs and dissolved WERs; variability in the concentration of nontoxic particulate metal will contribute to the variability in a total recoverable WER, but not to the variability in a dissolved WER. Thus, dissolved

WERs are expected to be less variable than total recoverable WERs, especially where events commonly increase TSS and/or TOC. In some cases, therefore, appropriate use of analytical chemistry can greatly increase the usefulness of the experimental determination of WERs. The concerns regarding variability are increased if an upstream effluent contributes to the WER.

When a WER is determined in downstream water, the four major sources of variability in the WER are (a) variability in the quality of the upstream water, which might be related to season and/or flow, (b) experimental variation, (c) variability in the composition of the effluent, and (d) variability in the ratio of the flows of the upstream water and the effluent. The considerations regarding the first two are the same as for WERs determined using upstream water; because of the additional sources of variability, WERs determined using downstream water are likely to be more variable than WERs determined using upstream water.

It would be desirable if a sufficient number of WERs could be determined to define the variable factors in the effluent and in the upstream water that contribute to the variability in WERs that are determined using downstream water. Not only is this likely to be very difficult in most cases, but it is also possible that the WER will be dependent on interactions between constituents of the effluent and the upstream water, i.e., the eWER and uWER might be additive, more-than-additive, or less-than-additive (see Appendix G). When interaction occurs, in order to completely understand the variability of WERs determined using downstream water, sufficient tests would have to be conducted to determine the means and variances of:

- a. the effluent component of the WER.
- b. the upstream component of the WER.
- c. any interaction between the two components.

An interaction might occur, for example, if the toxicity of a metal is affected by pH, and the pH and/or the buffering capacity of the effluent and/or the upstream water vary considerably.

An increase in the variability of WERs decreases the usefulness of any one WER. Compensation for this decrease in usefulness can be attempted by determining WERs at more times; although this will provide more data, it will not necessarily provide a proportionate increase in understanding. Rather than determining WERs at more times, a better use of resources might be to obtain more information concerning a smaller number of specially selected occasions.

It is likely that some cases will be so complex that achieving even a reasonable understanding will require unreasonable resources. In contrast, some WERs determined using the

methods presented herein might be relatively easy to understand if appropriate chemical measurements are performed when WERs are determined.

1. If the variation of the total recoverable WER is substantially greater than the variation of the comparable dissolved WER, there is probably a variable and substantial concentration of particulate nontoxic metal. It might be advantageous to use a dissolved WER just because it will have less variability than a total recoverable WER.
2. If the total recoverable and/or dissolved WER correlates with the total recoverable and/or dissolved concentration of metal in the site water, it is likely that a substantial percentage of the metal is nontoxic. In this case the WER will probably also depend on the concentration of effluent in the site water and on the concentration of metal in the effluent.

These approaches are more likely to be useful when WERs are determined using downstream water, rather than upstream water, unless both the magnitude of the WER and the concentration of the metal in the upstream water are elevated by an upstream effluent and/or events that increase TSS and/or TOC.

Both of these approaches can be applied to WERs that are determined using actual downstream water, but the second can probably provide much better information if it is used with WERs determined using simulated downstream water that is prepared by mixing a sample of the effluent with a sample of the upstream water. In this way the composition and characteristics of both the effluent and the upstream water can be determined, and the exact ratio in the downstream water is known.

Use of simulated downstream water is also a way to study the relation between the WER and the ratio of effluent to upstream water at one point in time, which is the most direct way to test for additivity of the eWER and the uWER (see Appendix G). This can be viewed as a test of the assumption that WERs determined using downstream water will decrease as the concentration of effluent decreases. If this assumption is true, as the flow increases, the concentration of effluent in the downstream water will decrease and the WER will decrease. Obtaining such information at one point in time is useful, but confirmation at one or more other times would be much more useful.

E. The fate of metal that has reduced or no toxicity.

Metal that has reduced or no toxicity at the end of the pipe might be more toxic at some time in the future. For example, metal that is in the water column and is not toxic now might become more toxic in the water column later or might move into

the sediment and become toxic. If a WER allows a surface water to contain as much toxic metal as is acceptable, the WER would not be adequately protective if metal that was nontoxic when the WER was determined became toxic in the water column, unless a compensating change occurred. Studies of the fate of metals need to address not only the changes that take place, but also the rates of the changes.

Concern about the fate of discharged metal justifiably raises concern about the possibility that metals might contaminate sediments. The possibility of contamination of sediment by toxic and/or nontoxic metal in the water column was one of the concerns that led to the establishment of EPA's sediment quality criteria program, which is developing guidelines and criteria to protect sediment. A separate program was necessary because ambient water quality criteria are not designed to protect sediment. Insofar as technology-based controls and water quality criteria reduce the discharge of metals, they tend to reduce the possibility of contamination of sediment. Conversely, insofar as WERs allow an increase in the discharge of metals, they tend to increase the possibility of contamination of sediment.

When WERs are determined in upstream water, the concern about the fate of metal with reduced or no toxicity is usually small because the WERs are usually small. In addition, the factors that result in upstream WERs being greater than 1.0 usually are (a) natural organic materials such as humic acids and (b) water quality characteristics such as hardness, alkalinity, and pH. It is easy to assume that natural organic materials will not degrade rapidly, and it is easy to monitor changes in hardness, alkalinity, and pH. Thus there is usually little concern about the fate of the metal when WERs are determined in upstream water, especially if the WER is small. If the WER is large and possibly due at least in part to an upstream effluent, there is more concern about the fate of metal that has reduced or no toxicity.

When WERs are determined in downstream water, effluents are allowed to contain virtually unlimited amounts of nontoxic particulate metal and nontoxic dissolved metal. It would seem prudent to obtain some data concerning whether the nontoxic metal might become toxic at some time in the future whenever (1) the concentration of nontoxic metal is large, (2) the concentration of dissolved metal is below the dissolved national criterion but the concentration of total recoverable metal is substantially above the total recoverable national criterion, or (3) the site-specific criterion is substantially above the national criterion. It would seem appropriate to:

- a. Generate some data concerning whether "fate" (i.e., environmental processes) will cause any of the nontoxic metal to become toxic due to oxidation of organic matter,

oxidation of sulfides, etc. For example, a WER could be determined using a sample of actual or simulated downstream water, the sample aerated for a period of time (e.g., two weeks), the pH adjusted if necessary, and another WER determined. If aeration reduced the WER, shorter and longer periods of aeration could be used to study the rate of change.

- b. Determine the effect of a change in water quality characteristics on the WER; for example, determine the effect of lowering the pH on the WER if influent lowers the pH of the downstream water within the area to which the site-specific criterion is to apply.
- c. Determine a WER in actual downstream water to demonstrate whether downstream conditions change sufficiently (possibly due to degradation of organic matter, multiple dischargers, etc.) to lower the WER more than the concentration of the metal is lowered.

If environmental processes cause nontoxic metal to become toxic, it is important to determine whether the time scale involves days, weeks, or years.

Summary

When WERs are determined using downstream water, the site water contains effluent and the WER will take into account not only the constituents of the upstream water, but also the toxic and nontoxic metal and other constituents of the effluent as they exist after mixing with upstream water. The determination of the WER automatically takes into account any additivity, synergism, or antagonism between the metal and components of the effluent and/or the upstream water. The effect of calcium, magnesium, and various heavy metals on competitive binding by such organic materials as humic acid is also taken into account. Therefore, a site-specific criterion derived using a WER is likely to be more appropriate for a site than a national, state, or recalculated criterion not only because it takes into account the water quality characteristics of the site water but also because it takes into account other constituents in the effluent and upstream water.

Determination of WERs using downstream water causes a general increase in the complexity, magnitude, and variability of WERs, and an increase in concern about the fate of metal that has reduced or no toxicity at the end of the pipe. In addition, there are some other drawbacks with the use of downstream water in the determination of a WER:

1. It might serve as a disincentive for some dischargers to remove any more organic carbon and/or particulate matter than required, although WERs for some metals will not be related to the concentration of TOC or TSS.

2. If conditions change, a WER might decrease in the future. This is not a problem if the decrease is due to a reduction in nontoxic metal, but it might be a problem if the decrease is due to a decrease in TOC or TSS or an increase in competitive binding.
3. If a WER is determined when the effluent contains refractory metal but a change in operations results in the discharge of toxic metal in place of refractory metal, the site-specific criterion and the permit limits will not provide adequate protection. In most cases chemical monitoring probably will not detect such a change, but toxicological monitoring probably will.

Use of WERs that are determined using downstream water rather than upstream water increases:

1. The importance of understanding the various issues involved in the determination and use of WERs.
2. The importance of obtaining data that will provide understanding rather than obtaining data that will result in the highest or lowest WER.
3. The appropriateness of site-specific criteria.
4. The resources needed to determine a WER.
5. The resources needed to use a WER.
6. The resources needed to monitor the acceptability of the downstream water.

A WER determined using upstream water will usually be smaller, less variable, and simpler to implement than a WER determined using downstream water. Although in some situations a downstream WER might be smaller than an upstream WER, the important consideration is that a WER should be determined using the water to which it is to apply.

References

U.S. EPA. 1983. Water Quality Standards Handbook. Office of Water Regulations and Standards, Washington, DC.

U.S. EPA. 1984. Guidelines for Deriving Numerical Aquatic Site-Specific Water Quality Criteria by Modifying National Criteria. EPA-600/3-84-099 or PB85-121101. National Technical Information Service, Springfield, VA.

U.S. EPA. 1992. Interim Guidance on Interpretation and Implementation of Aquatic Life Criteria for Metals. Office of Science and Technology, Health and Ecological Criteria Division, Washington, DC.

Appendix B: The Recalculation Procedure

NOTE: The National Toxics Rule (NTR) does not allow use of the Recalculation Procedure in the derivation of a site-specific criterion. Thus nothing in this appendix applies to jurisdictions that are subject to the NTR.

The Recalculation Procedure is intended to cause a site-specific criterion to appropriately differ from a national aquatic life criterion if justified by demonstrated pertinent toxicological differences between the aquatic species that occur at the site and those that were used in the derivation of the national criterion. There are at least three reasons why such differences might exist between the two sets of species. First, the national dataset contains aquatic species that are sensitive to many pollutants, but these and comparably sensitive species might not occur at the site. Second, a species that is critical at the site might be sensitive to the pollutant and require a lower criterion. (A critical species is a species that is commercially or recreationally important at the site, a species that exists at the site and is listed as threatened or endangered under section 4 of the Endangered Species Act, or a species for which there is evidence that the loss of the species from the site is likely to cause an unacceptable impact on a commercially or recreationally important species, a threatened or endangered species, the abundances of a variety of other species, or the structure or function of the community.) Third, the species that occur at the site might represent a narrower mix of species than those in the national dataset due to a limited range of natural environmental conditions. The procedure presented here is structured so that corrections and additions can be made to the national dataset without the deletion process being used to take into account taxa that do and do not occur at the site; in effect, this procedure makes it possible to update the national aquatic life criterion.

The phrase "occur at the site" includes the species, genera, families, orders, classes, and phyla that:

- a. are usually present at the site.
- b. are present at the site only seasonally due to migration.
- c. are present intermittently because they periodically return to or extend their ranges into the site.
- d. were present at the site in the past, are not currently present at the site due to degraded conditions, and are expected to return to the site when conditions improve.
- e. are present in nearby bodies of water, are not currently present at the site due to degraded conditions, and are expected to be present at the site when conditions improve.

The taxa that "occur at the site" cannot be determined merely by sampling downstream and/or upstream of the site at one point in time. "Occur at the site" does not include taxa that were once

present at the site but cannot exist at the site now due to permanent physical alteration of the habitat at the site resulting from dams, etc.

The definition of the "site" can be extremely important when using the Recalculation Procedure. For example, the number of taxa that occur at the site will generally decrease as the size of the site decreases. Also, if the site is defined to be very small, the permit limit might be controlled by a criterion that applies outside (e.g., downstream of) the site.

Note: If the variety of aquatic invertebrates, amphibians, and fishes is so limited that species in fewer than eight families occur at the site, the general Recalculation Procedure is not applicable and the following special version of the Recalculation Procedure **must** be used:

1. Data **must** be available for at least one species in each of the families that occur at the site.
2. The lowest Species Mean Acute Value that is available for a species that occurs at the site **must** be used as the FAV.
3. The site-specific CMC and CCC **must** be calculated as described below in part 2 of step E, which is titled "Determination of the CMC and/or CCC".

The concept of the Recalculation Procedure is to create a dataset that is appropriate for deriving a site-specific criterion by modifying the national dataset in some or all of three ways:

- a. Correction of data that are in the national dataset.
- b. Addition of data to the national dataset.
- c. Deletion of data that are in the national dataset.

All corrections and additions that have been approved by U.S. EPA are required, whereas use of the deletion process is optional. The Recalculation Procedure is more likely to result in lowering a criterion if the net result of addition and deletion is to decrease the number of genera in the dataset, whereas the procedure is more likely to result in raising a criterion if the net result of addition and deletion is to increase the number of genera in the dataset.

The Recalculation Procedure consists of the following steps:

- A. Corrections are made in the national dataset.
- B. Additions are made to the national dataset.
- C. The deletion process may be applied if desired.
- D. If the new dataset does not satisfy the applicable Minimum Data Requirements (MDRs), additional pertinent data **must** be generated; if the new data are approved by the U.S. EPA, the Recalculation Procedure **must** be started again at step B with the addition of the new data.
- E. The new CMC or CCC or both are determined.
- F. A report is written.

Each step is discussed in more detail below.

A. Corrections

1. Only corrections approved by the U.S. EPA may be made.
2. The concept of "correction" includes removal of data that should not have been in the national dataset in the first place. The concept of "correction" does not include removal of a datum from the national dataset just because the quality of the datum is claimed to be suspect. If additional data are available for the same species, the U.S. EPA will decide which data should be used, based on the available guidance (U.S. EPA 1985); also, data based on measured concentrations are usually preferable to those based on nominal concentrations.
3. Two kinds of corrections are possible:
 - a. The first includes those corrections that are known to and have been approved by the U.S. EPA; a list of these will be available from the U.S. EPA.
 - b. The second includes those corrections that are submitted to the U.S. EPA for approval. If approved, these will be added to EPA's list of approved corrections.
4. Selective corrections are not allowed. All corrections on EPA's newest list **must** be made.

B. Additions

1. Only additions approved by the U.S. EPA may be made.
2. Two kinds of additions are possible:
 - a. The first includes those additions that are known to and have been approved by the U.S. EPA; a list of these will be available from the U.S. EPA.
 - b. The second includes those additions that are submitted to the U.S. EPA for approval. If approved, these will be added to EPA's list of approved additions.
3. Selective additions are not allowed. All additions on EPA's newest list **must** be made.

C. The Deletion Process

The basic principles are:

1. Additions and corrections **must** be made as per steps A and B above, before the deletion process is performed.
2. Selective deletions are not allowed. If any species is to be deleted, the deletion process described below **must** be applied to all species in the national dataset, after any necessary corrections and additions have been made to the national dataset. The deletion process specifies which species **must** be deleted and which species **must not** be deleted. Use of the deletion process is optional, but no deletions are optional when the deletion process is used.
3. Comprehensive information **must** be available concerning what species occur at the site; a species cannot be deleted based

on incomplete information concerning the species that do and do not satisfy the definition of "occur at the site".

4. Data might have to be generated before the deletion process is begun:

- a. Acceptable pertinent toxicological data **must** be available for at least one species in each class of aquatic plants, invertebrates, amphibians, and fish that contains a species that is a critical species at the site.
- b. For each aquatic plant, invertebrate, amphibian, and fish species that occurs at the site and is listed as threatened or endangered under section 4 of the Endangered Species Act, data **must** be available or be generated for an acceptable surrogate species. Data for each surrogate species **must** be used as if they are data for species that occur at the site.

If additional data are generated using acceptable procedures (U.S. EPA 1985) and they are approved by the U.S. EPA, the Recalculation Procedure **must** be started again at step B with the addition of the new data.

5. Data might have to be generated after the deletion process is completed. Even if one or more species are deleted, there still are MDRs (see step D below) that **must** be satisfied. If the data remaining after deletion do not satisfy the applicable MDRs, additional toxicity tests **must** be conducted using acceptable procedures (U.S. EPA 1985) so that all MDRs are satisfied. If the new data are approved by the U.S. EPA, the Recalculation Procedure **must** be started again at step B with the addition of new data.
6. Chronic tests do not have to be conducted because the national Final Acute-Chronic Ratio (FACR) may be used in the derivation of the site-specific Final Chronic Value (FCV). If acute-chronic ratios (ACRs) are available or are generated so that the chronic MDRs are satisfied using only species that occur at the site, a site-specific FACR may be derived and used in place of the national FACR. Because a FACR was not used in the derivation of the freshwater CCC for cadmium, this CCC can only be modified the same way as a FAV; what is acceptable will depend on which species are deleted.

If any species are to be deleted, the following deletion process **must** be applied:

- a. Obtain a copy of the national dataset, i.e., tables 1, 2, and 3 in the national criteria document (see Appendix E).
- b. Make corrections in and/or additions to the national dataset as described in steps A and B above.
- c. Group all the species in the dataset taxonomically by phylum, class, order, family, genus, and species.
- d. Circle each species that satisfies the definition of "occur at the site" as presented on the first page of this appendix, and including any data for species that are surrogates of threatened or endangered species that occur at the site.

e. Use the following step-wise process to determine which of the uncircled species **must** be deleted and which **must not** be deleted:

1. Does the genus occur at the site?

If "No", go to step 2.

If "Yes", are there one or more species in the genus that occur at the site but are not in the dataset?

If "No", go to step 2.

If "Yes", retain the uncircled species.*

2. Does the family occur at the site?

If "No", go to step 3.

If "Yes", are there one or more genera in the family that occur at the site but are not in the dataset?

If "No", go to step 3.

If "Yes", retain the uncircled species.*

3. Does the order occur at the site?

If "No", go to step 4.

If "Yes", does the dataset contain a circled species that is in the same order?

If "No", retain the uncircled species.*

If "Yes", delete the uncircled species.*

4. Does the class occur at the site?

If "No", go to step 5.

If "Yes", does the dataset contain a circled species that is in the same class?

If "No", retain the uncircled species.*

If "Yes", delete the uncircled species.*

5. Does the phylum occur at the site?

If "No", delete the uncircled species.*

If "Yes", does the dataset contain a circled species that is in the same phylum?

If "No", retain the uncircled species.*

If "Yes", delete the uncircled species.*

* = Continue the deletion process by starting at step 1 for another uncircled species unless all uncircled species in the dataset have been considered.

The species that are circled and those that are retained constitute the site-specific dataset. (An example of the deletion process is given in Figure B1.)

This deletion process is designed to ensure that:

a. Each species that occurs both in the national dataset and at the site also occurs in the site-specific dataset.

- b. Each species that occurs at the site but does not occur in the national dataset is represented in the site-specific dataset by all species in the national dataset that are in the same genus.
- c. Each genus that occurs at the site but does not occur in the national dataset is represented in the site-specific dataset by all genera in the national dataset that are in the same family.
- d. Each order, class, and phylum that occurs both in the national dataset and at the site is represented in the site-specific dataset by the one or more species in the national dataset that are most closely related to a species that occurs at the site.

D. Checking the Minimum Data Requirements

The initial MDRs for the Recalculation Procedure are the same as those for the derivation of a national criterion. If a specific requirement cannot be satisfied after deletion because that kind of species does not occur at the site, a taxonomically similar species **must** be substituted in order to meet the eight MDRs:

If no species of the kind required occurs at the site, but a species in the same order does, the MDR can only be satisfied by data for a species that occurs at the site and is in that order; if no species in the order occurs at the site, but a species in the class does, the MDR can only be satisfied by data for a species that occurs at the site and is in that class. If no species in the same class occurs at the site, but a species in the phylum does, the MDR can only be satisfied by data for a species that occurs at the site and is in that phylum. If no species in the same phylum occurs at the site, any species that occurs at the site and is not used to satisfy a different MDR can be used to satisfy the MDR. If additional data are generated using acceptable procedures (U.S. EPA 1985) and they are approved by the U.S. EPA, the Recalculation Procedure **must** be started again at step B with the addition of the new data.

If fewer than eight families of aquatic invertebrates, amphibians, and fishes occur at the site, a Species Mean Acute Value **must** be available for at least one species in each of the families and the special version of the Recalculation Procedure described on the second page of this appendix **must** be used.

E. Determining the CMC and/or CCC

1. Determining the FAV:

- a. If the eight family MDRs are satisfied, the site-specific FAV **must** be calculated from Genus Mean Acute Values using

- the procedure described in the national aquatic life guidelines (U.S. EPA 1985).
- b. If fewer than eight families of aquatic invertebrates, amphibians, and fishes occur at the site, the lowest Species Mean Acute Value that is available for a species that occurs at the site **must** be used as the FAV, as per the special version of the Recalculation Procedure described on the second page of this appendix.
 2. The site-specific CMC **must** be calculated by dividing the site-specific FAV by 2. The site-specific FCV **must** be calculated by dividing the site-specific FAV by the national FACR (or by a site-specific FACR if one is derived). (Because a FACR was not used to derive the national CCC for cadmium in fresh water, the site-specific CCC equals the site-specific FCV.)
 3. The calculated FAV, CMC, and/or CCC **must** be lowered, if necessary, to (1) protect an aquatic plant, invertebrate, amphibian, or fish species that is a critical species at the site, and (2) ensure that the criterion is not likely to jeopardize the continued existence of any endangered or threatened species listed under section 4 of the Endangered Species Act or result in the destruction or adverse modification of such species' critical habitat.

F. Writing the Report

The report of the results of use of the Recalculation Procedure **must** include:

1. A list of all species of aquatic invertebrates, amphibians, and fishes that are known to "occur at the site", along with the source of the information.
2. A list of all aquatic plant, invertebrate, amphibian, and fish species that are critical species at the site, including all species that occur at the site and are listed as threatened or endangered under section 4 of the Endangered Species Act.
3. A site-specific version of Table 1 from a criteria document produced by the U.S. EPA after 1984.
4. A site-specific version of Table 3 from a criteria document produced by the U.S. EPA after 1984.
5. A list of all species that were deleted.
6. The new calculated FAV, CMC, and/or CCC.
7. The lowered FAV, CMC, and/or CCC, if one or more were lowered to protect a specific species.

Reference

U.S. EPA. 1985. Guidelines for Deriving Numerical National Water Quality Criteria for the Protection of Aquatic Organisms and Their Uses. PB85-227049. National Technical Information Service, Springfield, VA.

Figure B1: An Example of the Deletion Process Using Three Phyla

SPECIES THAT ARE IN THE THREE PHyla AND OCCUR AT THE SITE				
<u>Phylum</u>	<u>Class</u>	<u>Order</u>	<u>Family</u>	<u>Species</u>
Annelida	Hirudin.	Rhynchob.	Glossiph.	Glossip. complanata
Bryozoa	(No species in this phylum occur at the site.)			
Chordata	Osteich.	Cyprinif.	Cyprinid.	Carassius auratus
Chordata	Osteich.	Cyprinif.	Cyprinid.	Notropis anogenus
Chordata	Osteich.	Cyprinif.	Cyprinid.	Phoxinus eos
Chordata	Osteich.	Cyprinif.	Catostom.	Carpiodes carpio
Chordata	Osteich.	Salmonif.	Osmerida.	Osmerus mordax
Chordata	Osteich.	Percifor.	Centrarc.	Lepomis cyanellus
Chordata	Osteich.	Percifor.	Centrarc.	Lepomis humilis
Chordata	Amphibia	Caudata	Ambystom.	Ambystoma gracile

SPECIES THAT ARE IN THE THREE PHyla AND IN THE NATIONAL DATASET					
<u>Phylum</u>	<u>Class</u>	<u>Order</u>	<u>Family</u>	<u>Species</u>	<u>Code</u>
Annelida	Oligoch.	Haplotax.	Tubifici.	Tubifex tubifex	P
Bryozoa	Phylact.	---	Lophopod.	Lophopod. carteri	D
Chordata	Cephala.	Petromyz.	Petromyz.	Petromyzon marinus	D
Chordata	Osteich.	Cyprinif.	Cyprinid.	Carassius auratus	S
Chordata	Osteich.	Cyprinif.	Cyprinid.	Notropis hudsonius	G
Chordata	Osteich.	Cyprinif.	Cyprinid.	Notropis stramineus	G
Chordata	Osteich.	Cyprinif.	Cyprinid.	Phoxinus eos	S
Chordata	Osteich.	Cyprinif.	Cyprinid.	Phoxinus oreas	D
Chordata	Osteich.	Cyprinif.	Cyprinid.	Tinca tinca	D
Chordata	Osteich.	Cyprinif.	Catostom.	Ictiobus bubalus	F
Chordata	Osteich.	Salmonif.	Salmonid.	Oncorhynchus mykiss	O
Chordata	Osteich.	Percifor.	Centrarc.	Lepomis cyanellus	S
Chordata	Osteich.	Percifor.	Centrarc.	Lepomis macrochirus	G
Chordata	Osteich.	Percifor.	Percidae	Perca flavescens	D
Chordata	Amphibia	Anura	Pipidae	Xenopus laevis	C

Explanations of Codes:

- S = retained because this Species occurs at the site.
- G = retained because there is a species in this Genus that occurs at the site but not in the national dataset.
- F = retained because there is a genus in this Family that occurs at the site but not in the national dataset.
- O = retained because this Order occurs at the site and is not represented by a lower taxon.
- C = retained because this Class occurs at the site and is not represented by a lower taxon.
- P = retained because this Phylum occurs at the site and is not represented by a lower taxon.
- D = deleted because this species does not satisfy any of the requirements for retaining species.

Appendix C: Guidance Concerning the Use of "Clean Techniques" and QA/QC when Measuring Trace Metals

Note: This version of this appendix contains more information than the version that was Appendix B of Prothro (1993).

Recent information (Shiller and Boyle 1987; Windom et al. 1991) has raised questions concerning the quality of reported concentrations of trace metals in both fresh and salt (estuarine and marine) surface waters. A lack of awareness of true ambient concentrations of metals in fresh and salt surface waters can be both a cause and a result of the problem. The ranges of dissolved metals that are typical in surface waters of the United States away from the immediate influence of discharges (Bruland 1983; Shiller and Boyle 1985, 1987; Trefry et al. 1986; Windom et al. 1991) are:

Metal	Salt water ($\mu\text{g/L}$)	Fresh water ($\mu\text{g/L}$)
Cadmium	0.01 to 0.2	0.002 to 0.08
Copper	0.1 to 3.	0.4 to 4.
Lead	0.01 to 1.	0.01 to 0.19
Nickel	0.3 to 5.	1. to 2.
Silver	0.005 to 0.2	-----
Zinc	0.1 to 15.	0.03 to 5.

The U.S. EPA (1983, 1991) has published analytical methods for monitoring metals in waters and wastewaters, but these methods are inadequate for determination of ambient concentrations of some metals in some surface waters. Accurate and precise measurement of these low concentrations requires appropriate attention to seven areas:

1. Use of "clean techniques" during collecting, handling, storing, preparing, and analyzing samples to avoid contamination.
2. Use of analytical methods that have sufficiently low detection limits.
3. Avoidance of interference in the quantification (instrumental analysis) step.
4. Use of blanks to assess contamination.
5. Use of matrix spikes (sample spikes) and certified reference materials (CRMs) to assess interference and contamination.
6. Use of replicates to assess precision.
7. Use of certified standards.

In a strict sense, the term "clean techniques" refers to techniques that reduce contamination and enable the accurate and precise measurement of trace metals in fresh and salt surface waters. In a broader sense, the term also refers to related issues concerning detection limits, quality control, and quality

assurance. Documenting data quality demonstrates the amount of confidence that can be placed in the data, whereas increasing the sensitivity of methods reduces the problem of deciding how to interpret results that are reported to be below detection limits.

This appendix is written for those analytical laboratories that want guidance concerning ways to lower detection limits, increase accuracy, and/or increase precision. The ways to achieve these goals are to increase the sensitivity of the analytical methods, decrease contamination, and decrease interference. Ideally, validation of a procedure for measuring concentrations of metals in surface water requires demonstration that agreement can be obtained using completely different procedures beginning with the sampling step and continuing through the quantification step (Bruland et al. 1979), but few laboratories have the resources to compare two different procedures. Laboratories can, however, (a) use techniques that others have found useful for improving detection limits, accuracy, and precision, and (b) document data quality through use of blanks, spikes, CRMs, replicates, and standards.

Nothing contained or not contained in this appendix adds to or subtracts from any regulatory requirement set forth in other EPA documents concerning analyses of metals. A WER can be acceptably determined without the use of clean techniques as long as the detection limits, accuracy, and precision are acceptable. No QA/QC requirements beyond those that apply to measuring metals in effluents are necessary for the determination of WERs. The word "must" is not used in this appendix. Some items, however, are considered so important by analytical chemists who have worked to increase accuracy and precision and lower detection limits in trace-metal analysis that "**should**" is in bold print to draw attention to the item. Most such items are emphasized because they have been found to have received inadequate attention in some laboratories performing trace-metal analyses.

In general, in order to achieve accurate and precise measurement of a particular concentration, both the detection limit and the blanks should be less than one-tenth of that concentration. Therefore, the term "metal-free" can be interpreted to mean that the total amount of contamination that occurs during sample collection and processing (e.g., from gloves, sample containers, labware, sampling apparatus, cleaning solutions, air, reagents, etc.) is sufficiently low that blanks are less than one-tenth of the lowest concentration that needs to be measured.

Atmospheric particulates can be a major source of contamination (Moody 1982; Adeloju and Bond 1985). The term "class-100" refers to a specification concerning the amount of particulates in air (Moody 1982); although the specification says nothing about the composition of the particulates, generic control of particulates can greatly reduce trace-metal blanks. Except during collection

of samples, initial cleaning of equipment, and handling of samples containing high concentrations of metals, all handling of samples, sample containers, labware, and sampling apparatus should be performed in a class-100 bench, room, or glove box.

Neither the "ultraclean techniques" that might be necessary when trace analyses of mercury are performed nor safety in analytical laboratories is addressed herein. Other documents should be consulted if one or both of these topics are of concern.

Avoiding contamination by use of "clean techniques"

Measurement of trace metals in surface waters should take into account the potential for contamination during each step in the process. Regardless of the specific procedures used for collection, handling, storage, preparation (digestion, filtration, and/or extraction), and quantification (instrumental analysis), the general principles of contamination control should be applied. Some specific recommendations are:

- a. Powder-free (non-talc, class-100) latex, polyethylene, or polyvinyl chloride (PVC, vinyl) gloves **should** be worn during all steps from sample collection to analysis. (Talc seems to be a particular problem with zinc; gloves made with talc cannot be decontaminated sufficiently.) Gloves should only contact surfaces that are metal-free; gloves should be changed if even suspected of contamination.
- b. The acid used to acidify samples for preservation and digestion and to acidify water for final cleaning of labware, sampling apparatus, and sample containers **should** be metal-free. The quality of the acid used should be better than reagent-grade. Each lot of acid **should** be analyzed for the metal(s) of interest before use.
- c. The water used to prepare acidic cleaning solutions and to rinse labware, sample containers, and sampling apparatus may be prepared by distillation, deionization, or reverse osmosis, and **should** be demonstrated to be metal-free.
- d. The work area, including bench tops and hoods, should be cleaned (e.g., washed and wiped dry with lint-free, class-100 wipes) frequently to remove contamination.
- e. All handling of samples in the laboratory, including filtering and analysis, **should** be performed in a class-100 clean bench or a glove box fed by particle-free air or nitrogen; ideally the clean bench or glove box should be located within a class-100 clean room.
- f. Labware, reagents, sampling apparatus, and sample containers **should** never be left open to the atmosphere; they should be stored in a class-100 bench, covered with plastic wrap, stored in a plastic box, or turned upside down on a clean surface. Minimizing the time between cleaning and using will help minimize contamination.

- g. Separate sets of sample containers, labware, and sampling apparatus should be dedicated for different kinds of samples, e.g., surface water samples, effluent samples, etc.
- h. To avoid contamination of clean rooms, samples that contain very high concentrations of metals and do not require use of "clean techniques" **should not** be brought into clean rooms.
- i. Acid-cleaned plastic, such as high-density polyethylene (HDPE), low-density polyethylene (LDPE), or a fluoroplastic, **should** be the only material that ever contacts a sample, except possibly during digestion for the total recoverable measurement.
 - 1. Total recoverable samples can be digested in some plastic containers.
 - 2. HDPE and LDPE might not be acceptable for mercury.
 - 3. Even if acidified, samples and standards containing silver should be in amber containers.
- j. All labware, sample containers, and sampling apparatus **should** be acid-cleaned before use or reuse.
 - 1. Sample containers, sampling apparatus, tubing, membrane filters, filter assemblies, and other labware **should** be soaked in acid until metal-free. The amount of cleaning necessary might depend on the amount of contamination and the length of time the item will be in contact with samples. For example, if an acidified sample will be stored in a sample container for three weeks, ideally the container should have been soaked in an acidified metal-free solution for at least three weeks.
 - 2. It might be desirable to perform initial cleaning, for which reagent-grade acid may be used, before the items are taken into a clean room. For most metals, items should be either (a) soaked in 10 percent concentrated nitric acid at 50°C for at least one hour, or (b) soaked in 50 percent concentrated nitric acid at room temperature for at least two days; for arsenic and mercury, soaking for up to two weeks at 50°C in 10 percent concentrated nitric acid might be required. For plastics that might be damaged by strong nitric acid, such as polycarbonate and possibly HDPE and LDPE, soaking in 10 percent concentrated hydrochloric acid, either in place of or before soaking in a nitric acid solution, might be desirable.
 - 3. Chromic acid **should not** be used to clean items that will be used in analysis of metals.
 - 4. Final soaking and cleaning of sample containers, labware, and sampling apparatus **should** be performed in a class-100 clean room using metal-free acid and water. The solution in an acid bath **should** be analyzed periodically to demonstrate that it is metal-free.
- k. Labware, sampling apparatus, and sample containers **should** be stored appropriately after cleaning:
 - 1. After the labware and sampling apparatus are cleaned, they may be stored in a clean room in a weak acid bath prepared using metal-free acid and water. Before use, the items

- should be rinsed at least three times with metal-free water. After the final rinse, the items should be moved immediately, with the open end pointed down, to a class-100 clean bench. Items may be dried on a class-100 clean bench; items **should not** be dried in an oven or with laboratory towels. The sampling apparatus should be assembled in a class-100 clean room or bench and double-bagged in metal-free polyethylene zip-type bags for transport to the field; new bags are usually metal-free.
2. After sample containers are cleaned, they should be filled with metal-free water that has been acidified to a pH of 2 with metal-free nitric acid (about 0.5 mL per liter) for storage until use.
 1. Labware, sampling apparatus, and sample containers **should** be rinsed and not rinsed with sample as necessary to prevent high and low bias of analytical results because acid-cleaned plastic will sorb some metals from unacidified solutions.
 1. Because samples for the dissolved measurement are not acidified until after filtration, all sampling apparatus, sample containers, labware, filter holders, membrane filters, etc., that contact the sample before or during filtration **should** be rinsed with a portion of the solution and then that portion discarded.
 2. For the total recoverable measurement, labware, etc., that contact the sample only before it is acidified **should** be rinsed with sample, whereas items that contact the sample after it is acidified **should not** be rinsed. For example, the sampling apparatus should be rinsed because the sample will not be acidified until it is in a sample container, but the sample container should not be rinsed if the sample will be acidified in the sample container.
 3. If the total recoverable and dissolved measurements are to be performed on the same sample (rather than on two samples obtained at the same time and place), all the apparatus and labware, including the sample container, should be rinsed before the sample is placed in the sample container; then an unacidified aliquot should be removed for the total recoverable measurement (and acidified, digested, etc.) and an unacidified aliquot should be removed for the dissolved measurement (and filtered, acidified, etc.) (If a container is rinsed and filled with sample and an unacidified aliquot is removed for the dissolved measurement and then the solution in the container is acidified before removal of an aliquot for the total recoverable measurement, the resulting measured total recoverable concentration might be biased high because the acidification might desorb metal that had been sorbed onto the walls of the sample container; the amount of bias will depend on the relative volumes involved and on the amount of sorption and desorption.)
 - m. Field samples **should** be collected in a manner that eliminates the potential for contamination from sampling platforms,

- probes, etc. Exhaust from boats and the direction of wind and water currents should be taken into account. The people who collect the samples **should** be specifically trained on how to collect field samples. After collection, all handling of samples in the field that will expose the sample to air **should** be performed in a portable class-100 clean bench or glove box.
- n. Samples **should** be acidified (after filtration if dissolved metal is to be measured) to a pH of less than 2, except that the pH **should** be less than 1 for mercury. Acidification should be done in a clean room or bench, and so it might be desirable to wait and acidify samples in a laboratory rather than in the field. If samples are acidified in the field, metal-free acid can be transported in plastic bottles and poured into a plastic container from which acid can be removed and added to samples using plastic pipettes. Alternatively, plastic automatic dispensers can be used.
 - o. Such things as probes and thermometers **should not** be put in samples that are to be analyzed for metals. In particular, pH electrodes and mercury-in-glass thermometers **should not** be used if mercury is to be measured. If pH is measured, it **should** be done on a separate aliquot.
 - p. Sample handling should be minimized. For example, instead of pouring a sample into a graduated cylinder to measure the volume, the sample can be weighed after being poured into a tared container, which is less likely to be subject to error than weighing the container from which the sample is poured. (For saltwater samples, the salinity or density should be taken into account if weight is converted to volume.)
 - q. Each reagent used **should** be verified to be metal-free. If metal-free reagents are not commercially available, removal of metals will probably be necessary.
 - r. For the total recoverable measurement, samples should be digested in a class-100 bench, not in a metallic hood. If feasible, digestion should be done in the sample container by acidification and heating.
 - s. The longer the time between collection and analysis of samples, the greater the chance of contamination, loss, etc.
 - t. Samples should be stored in the dark, preferably between 0 and 4°C with no air space in the sample container.

Achieving low detection limits

- a. Extraction of the metal from the sample can be extremely useful if it simultaneously concentrates the metal and eliminates potential matrix interferences. For example, ammonium 1-pyrrolidinedithiocarbamate and/or diethylammonium diethyldithiocarbamate can extract cadmium, copper, lead, nickel, and zinc (Bruland et al. 1979; Nriagu et al. 1993).
- b. The detection limit should be less than ten percent of the lowest concentration that is to be measured.

Avoiding interferences

- a. Potential interferences **should** be assessed for the specific instrumental analysis technique used and for each metal to be measured.
- b. If direct analysis is used, the salt present in high-salinity saltwater samples is likely to cause interference in most instrumental techniques.
- c. As stated above, extraction of the metal from the sample is particularly useful because it simultaneously concentrates the metal and eliminates potential matrix interferences.

Using blanks to assess contamination

- a. A laboratory (procedural, method) blank consists of filling a sample container with analyzed metal-free water and processing (filtering, acidifying, etc.) the water through the laboratory procedure in exactly the same way as a sample. A laboratory blank **should** be included in each set of ten or fewer samples to check for contamination in the laboratory, and **should** contain less than ten percent of the lowest concentration that is to be measured. Separate laboratory blanks **should** be processed for the total recoverable and dissolved measurements, if both measurements are performed.
- b. A field (trip) blank consists of filling a sample container with analyzed metal-free water in the laboratory, taking the container to the site, processing the water through tubing, filter, etc., collecting the water in a sample container, and acidifying the water the same as a field sample. A field blank **should** be processed for each sampling trip. Separate field blanks **should** be processed for the total recoverable measurement and for the dissolved measurement, if filtrations are performed at the site. Field blanks **should** be processed in the laboratory the same as laboratory blanks.

Assessing accuracy

- a. A calibration curve **should** be determined for each analytical run and the calibration should be checked about every tenth sample. Calibration solutions **should** be traceable back to a certified standard from the U.S. EPA or the National Institute of Science and Technology (NIST).
- b. A blind standard or a blind calibration solution **should** be included in each group of about twenty samples.
- c. At least one of the following **should** be included in each group of about twenty samples:
 1. A matrix spike (spiked sample; the method of known additions).

2. A CRM, if one is available in a matrix that closely approximates that of the samples. Values obtained for the CRM **should** be within the published values. The concentrations in blind standards and solutions, spikes, and CRMs **should not** be more than 5 times the median concentration expected to be present in the samples.

Assessing precision

- a. A sampling replicate **should** be included with each set of samples collected at each sampling location.
- b. If the volume of the sample is large enough, replicate analysis of at least one sample **should** be performed along with each group of about ten samples.

Special considerations concerning the dissolved measurement

Whereas total recoverable measurements are especially subject to contamination during digestion, dissolved measurements are subject to both loss and contamination during filtration.

- a. Because acid-cleaned plastic sorbs metal from unacidified solutions and because samples for the dissolved measurement are not acidified before filtration, all sampling apparatus, sample containers, labware, filter holders, and membrane filters that contact the sample before or during filtration **should** be conditioned by rinsing with a portion of the solution and discarding that portion.
- b. Filtrations **should** be performed using acid-cleaned plastic filter holders and acid-cleaned membrane filters. Samples **should not** be filtered through glass fiber filters, even if the filters have been cleaned with acid. If positive-pressure filtration is used, the air or gas **should** be passed through a 0.2- μ m in-line filter; if vacuum filtration is used, it **should** be performed on a class-100 bench.
- c. Plastic filter holders **should** be rinsed and/or dipped between filtrations, but they do not have to be soaked between filtrations if all the samples contain about the same concentrations of metal. It is best to filter samples from low to high concentrations. A membrane filter **should not** be used for more than one filtration. After each filtration, the membrane filter **should** be removed and discarded, and the filter holder **should** be either rinsed with metal-free water or dilute acid and dipped in a metal-free acid bath or rinsed at least twice with metal-free dilute acid; finally, the filter holder **should** be rinsed at least twice with metal-free water.
- d. For each sample to be filtered, the filter holder and membrane filter **should** be conditioned with the sample, i.e., an initial portion of the sample **should** be filtered and discarded.

The accuracy and precision of the dissolved measurement **should** be assessed periodically. A large volume of a buffered solution (such as aerated 0.05 N sodium bicarbonate for analyses in fresh water and a combination of sodium bicarbonate and sodium chloride for analyses in salt water) should be spiked so that the concentration of the metal of interest is in the range of the low concentrations that are to be measured. Sufficient samples should be taken alternately for (a) acidification in the same way as after filtration in the dissolved method and (b) filtration and acidification using the procedures specified in the dissolved method until ten samples have been processed in each way. The concentration of metal in each of the twenty samples should then be determined using the same analytical procedure. The means of the two groups of ten measurements should be within 10 percent, and the coefficient of variation for each group of ten should be less than 20 percent. Any values deleted as outliers **should** be acknowledged.

Reporting results

To indicate the quality of the data, reports of results of measurements of the concentrations of metals **should** include a description of the blanks, spikes, CRMs, replicates, and standards that were run, the number run, and the results obtained. All values deleted as outliers **should** be acknowledged.

Additional information

The items presented above are some of the important aspects of "clean techniques"; some aspects of quality assurance and quality control are also presented. This is not a definitive treatment of these topics; additional information that might be useful is available in such publications as Patterson and Settle (1976), Zief and Mitchell (1976), Bruland et al. (1979), Moody and Beary (1982), Moody (1982), Bruland (1983), Adeloju and Bond (1985), Berman and Yeats (1985), Byrd and Andreae (1986), Taylor (1987), Sakamoto-Arnold (1987), Tramontano et al. (1987), Puls and Barcelona (1989), Windom et al. (1991), U.S. EPA (1992), Horowitz et al. (1992), and Nriagu et al. (1993).

References

- Adeloju, S.B., and A.M. Bond. 1985. Influence of Laboratory Environment on the Precision and Accuracy of Trace Element Analysis. *Anal. Chem.* 57:1728-1733.
- Berman, S.S., and P.A. Yeats. 1985. Sampling of Seawater for Trace Metals. *CRC Reviews in Analytical Chemistry* 16:1-14.
- Bruland, K.W., R.P. Franks, G.A. Knauer, and J.H. Martin. 1979. Sampling and Analytical Methods for the Determination of Copper, Cadmium, Zinc, and Nickel at the Nanogram per Liter Level in Sea Water. *Anal. Chim. Acta* 105:233-245.
- Bruland, K.W. 1983. Trace Elements in Sea-water. In: *Chemical Oceanography*, Vol. 8. (J.P. Riley and R. Chester, eds.) Academic Press, New York, NY. pp. 157-220.
- Byrd, J.T., and M.O. Andreae. 1986. Dissolved and Particulate Tin in North Atlantic Seawater. *Marine Chem.* 19:193-200.
- Horowitz, A.J., K.A. Elrick, and M.R. Colberg. 1992. The Effect of Membrane Filtration Artifacts on Dissolved Trace Element Concentrations. *Water Res.* 26:753-763.
- Moody, J.R. 1982. NBS Clean Laboratories for Trace Element Analysis. *Anal. Chem.* 54:1358A-1376A.
- Moody, J.R., and E.S. Beary. 1982. Purified Reagents for Trace Metal Analysis. *Talanta* 29:1003-1010.
- Nriagu, J.O., G. Lawson, H.K.T. Wong, and J.M. Azcue. 1993. A Protocol for Minimizing Contamination in the Analysis of Trace Metals in Great Lakes Waters. *J. Great Lakes Res.* 19:175-182.
- Patterson, C.C., and D.M. Settle. 1976. The Reduction in Orders of Magnitude Errors in Lead Analysis of Biological Materials and Natural Waters by Evaluating and Controlling the Extent and Sources of Industrial Lead Contamination Introduced during Sample Collection and Processing. In: *Accuracy in Trace Analysis: Sampling, Sample Handling, Analysis.* (P.D. LaFleur, ed.) National Bureau of Standards Spec. Publ. 422, U.S. Government Printing Office, Washington, DC.
- Prothro, M.G. 1993. Memorandum titled "Office of Water Policy and Technical Guidance on Interpretation and Implementation of Aquatic Life Metals Criteria". October 1.
- Puls, R.W., and M.J. Barcelona. 1989. Ground Water Sampling for Metals Analyses. EPA/540/4-89/001. National Technical Information Service, Springfield, VA.

- Sakamoto-Arnold, C.M., A.K. Hanson, Jr., D.L. Huizenga, and D.R. Kester. 1987. Spatial and Temporal Variability of Cadmium in Gulf Stream Warm-core Rings and Associated Waters. *J. Mar. Res.* 45:201-230.
- Shiller, A.M., and E. Boyle. 1985. Dissolved Zinc in Rivers. *Nature* 317:49-52.
- Shiller, A.M., and E.A. Boyle. 1987. Variability of Dissolved Trace Metals in the Mississippi River. *Geochim. Cosmochim. Acta* 51:3273-3277.
- Taylor, J.K. 1987. Quality Assurance of Chemical Measurements. Lewis Publishers, Chelsea, MI.
- Tramontano, J.M., J.R. Scudlark, and T.M. Church. 1987. A Method for the Collection, Handling, and Analysis of Trace Metals in Precipitation. *Environ. Sci. Technol.* 21:749-753.
- Trefry, J.H., T.A. Nelsen, R.P. Trocine, S. Metz., and T.W. Vetter. 1986. Trace Metal Fluxes through the Mississippi River Delta System. *Rapp. P.-v. Reun. Cons. int. Explor. Mer.* 186:277-288.
- U.S. EPA. 1983. Methods for Chemical Analysis of Water and Wastes. EPA-600/4-79-020. National Technical Information Service, Springfield, VA. Sections 4.1.1, 4.1.3, and 4.1.4
- U.S. EPA. 1991. Methods for the Determination of Metals in Environmental Samples. EPA-600/4-91-010. National Technical Information Service, Springfield, VA.
- U.S. EPA. 1992. Evaluation of Trace-Metal Levels in Ambient Waters and Tributaries to New York/New Jersey Harbor for Waste Load Allocation. Prepared by Battelle Ocean Sciences under Contract No. 68-C8-0105.
- Windom, H.L., J.T. Byrd, R.G. Smith, and F. Huan. 1991. Inadequacy of NASQAN Data for Assessing Metals Trends in the Nation's Rivers. *Environ. Sci. Technol.* 25:1137-1142. (Also see the comment and response: *Environ. Sci. Technol.* 25:1940-1941.)
- Zief, M., and J.W. Mitchell. 1976. Contamination Control in Trace Element Analysis. Chemical Analysis Series, Vol. 47. Wiley, New York, NY.

Appendix D: Relationships between WERs and the Chemistry and Toxicology of Metals

The aquatic toxicology of metals is complex in part because the chemistry of metals in water is complex. Metals usually exist in surface water in various combinations of particulate and dissolved forms, some of which are toxic and some of which are nontoxic. In addition, all toxic forms of a metal are not necessarily equally toxic, and various water quality characteristics can affect the relative concentrations and/or toxicities of some of the forms.

The toxicity of a metal has sometimes been reported to be proportional to the concentration or activity of a specific species of the metal. For example, Allen and Hansen (1993) summarized reports by several investigators that the toxicity of copper is related to the free cupric ion, but other data do not support a correlation (Erickson 1993a). For example, Borgmann (1983), Chapman and McCrady (1977), and French and Hunt (1986) found that toxicity expressed on the basis of cupric ion activity varied greatly with pH, and Cowan et al. (1986) concluded that at least one of the copper hydroxide species is toxic. Further, chloride and sulfate salts of calcium, magnesium, potassium, and sodium affect the toxicity of the cupric ion (Nelson et al. 1986). Similarly for aluminum, Wilkinson et al. (1993) concluded that "mortality was best predicted not by the free Al^{3+} activity but rather as a function of the sum $\Sigma([\text{Al}^{3+}] + [\text{AlF}^{2+}])$ " and that "no longer can the reduction of Al toxicity in the presence of organic acids be interpreted simply as a consequence of the decrease in the free Al^{3+} concentration".

Until a model has been demonstrated to explain the quantitative relationship between chemical and toxicological measurements, aquatic life criteria should be established in an environmentally conservative manner with provision for site-specific adjustment. Criteria should be expressed in terms of feasible analytical measurements that provide the necessary conservatism without substantially increasing the cost of implementation and site-specific adjustment. Thus current aquatic life criteria for metals are expressed in terms of the total recoverable measurement and/or the dissolved measurement, rather than a measurement that would be more difficult to perform and would still require empirical adjustment. The WER is operationally defined in terms of chemical and toxicological measurements to allow site-specific adjustments that account for differences between the toxicity of a metal in laboratory dilution water and in site water.

Forms of Metals

Even if the relationship of toxicity to the forms of metals is not understood well enough to allow setting site-specific water quality criteria without using empirical adjustments, appropriate use and interpretation of WERs requires an understanding of how changes in the relative concentrations of different forms of a metal might affect toxicity. Because WERs are defined on the basis of relationships between measurements of toxicity and measurements of total recoverable and/or dissolved metal, the toxicologically relevant distinction is between the forms of the metal that are toxic and nontoxic whereas the chemically relevant distinction is between the forms that are dissolved and particulate. "Dissolved metal" is defined here as "metal that passes through either a 0.45- μ m or a 0.40- μ m membrane filter" and "particulate metal" is defined as "total recoverable metal minus dissolved metal". Metal that is in or on particles that pass through the filter is operationally defined as "dissolved".

In addition, some species of metal can be converted from one form to another. Some conversions are the result of reequilibration in response to changes in water quality characteristics whereas others are due to such fate processes as oxidation of sulfides and/or organic matter. Reequilibration usually occurs faster than fate processes and probably results in any rapid changes that are due to effluent mixing with receiving water or changes in pH at a gill surface. To account for rapid changes due to reequilibration, the terms "labile" and "refractory" will be used herein to denote metal species that do and do not readily convert to other species when in a nonequilibrium condition, with "readily" referring to substantial progression toward equilibrium in less than about an hour. Although the toxicity and lability of a form of a metal are not merely yes/no properties, but rather involve gradations, a simple classification scheme such as this should be sufficient to establish the principles regarding how WERs are related to various operationally defined forms of metal and how this affects the determination and use of WERs.

Figure D1 presents the classification scheme that results from distinguishing forms of metal based on analytical methodology, toxicity tests, and lability, as described above. Metal that is not measured by the total recoverable measurement is assumed to be sufficiently nontoxic and refractory that it will not be further considered here. Allowance is made for toxicity due to particulate metal because some data indicate that particulate metal might contribute to toxicity and bioaccumulation, although other data imply that little or no toxicity can be ascribed to particulate metal (Erickson 1993b). Even if the toxicity of particulate metal is not negligible in a particular situation, a dissolved criterion will not be underprotective if the dissolved criterion was derived using a dissolved WER (see below) or if there are sufficient compensating factors.

Figure D1: A Scheme for Classifying Forms of Metal in Water

Total recoverable metal
 Dissolved
 Nontoxic
 Labile
 Refractory
 Toxic
 Labile
 Particulate
 Nontoxic
 Labile
 Refractory
 Toxic
 Labile
Metal not measured by the total recoverable measurement

Not only can some changes in water quality characteristics shift the relative concentrations of toxic and nontoxic labile species of a metal, some changes in water quality can also increase or decrease the toxicities of the toxic species of a metal and/or the sensitivities of aquatic organisms. Such changes might be caused by (a) a change in ionic strength that affects the activity of toxic species of the metal in water, (b) a physiological effect whereby an ion affects the permeability of a membrane and thereby alters both uptake and apparent toxicity, and (c) toxicological additivity, synergism, or antagonism due to effects within the organism.

Another possible complication is that a form of metal that is toxic to one aquatic organism might not be toxic to another. Although such differences between organisms have not been demonstrated, the possibility cannot be ruled out.

The Importance of Lability

The only common metal measurement that can be validly extrapolated from the effluent and the upstream water to the downstream water merely by taking dilution into account is the total recoverable measurement. A major reason this measurement is so useful is because it is the only measurement that obeys the law of mass balance (i.e., it is the only measurement that is conservative). Other metal measurements usually do not obey the law of mass balance because they measure some, but not all, of the labile species of metals. A measurement of refractory metal

would be conservative in terms of changes in water quality characteristics, but not necessarily in regards to fate processes; such a measurement has not been developed, however.

Permit limits apply to effluents, whereas water quality criteria apply to surface waters. If permit limits and water quality criteria are both expressed in terms of total recoverable metal, extrapolations from effluent to surface water only need to take dilution into account and can be performed as mass balance calculations. If either permit limits or water quality criteria or both are expressed in terms of any other metal measurement, lability needs to be taken into account, even if both are expressed in terms of the same measurement.

Extrapolations concerning labile species of metals from effluent to surface water depend to a large extent on the differences between the water quality characteristics of the effluent and those of the surface water. Although equilibrium models of the speciation of metals can provide insight, the interactions are too complex to be able to make useful nonempirical extrapolations from a wide variety of effluents to a wide variety of surface waters of either (a) the speciation of the metal or (b) a metal measurement other than total recoverable.

Empirical extrapolations can be performed fairly easily and the most common case will probably occur when permit limits are based on the total recoverable measurement but water quality criteria are based on the dissolved measurement. The empirical extrapolation is intended to answer the question "What percent of the total recoverable metal in the effluent becomes dissolved in the downstream water?" This question can be answered by:

- a. Collecting samples of effluent and upstream water.
- b. Measuring total recoverable metal and dissolved metal in both samples.
- c. Combining aliquots of the two samples in the ratio of the flows when the samples were obtained and mixing for an appropriate period of time under appropriate conditions.
- d. Measuring total recoverable metal and dissolved metal in the mixture.

An example is presented in Figure D2. This percentage cannot be extrapolated from one metal to another or from one effluent to another. The data needed to calculate the percentage will be obtained each time a WER is determined using simulated downstream water if both dissolved and total recoverable metal are measured in the effluent, upstream water, and simulated downstream water.

The interpretation of the percentage is not necessarily as straightforward as might be assumed. For example, some of the metal that is dissolved in the upstream water might sorb onto particulate matter in the effluent, which can be viewed as a detoxification of the upstream water by the effluent. Regardless of the interpretation, the described procedure provides a simple

way of relating the total recoverable concentration in the effluent to the concentration of concern in the downstream water. Because this empirical extrapolation can be used with any analytical measurement that is chosen as the basis for expression of aquatic life criteria, use of the total recoverable measurement to express permit limits on effluents does not place any restrictions on which analytical measurement can be used to express criteria. Further, even if both criteria and permit limits are expressed in terms of a measurement such as dissolved metal, an empirical extrapolation would still be necessary because dissolved metal is not likely to be conservative from effluent to downstream water.

Merits of Total Recoverable and Dissolved WERs and Criteria

A WER is operationally defined as the value of an endpoint obtained with a toxicity test using site water divided by the value of the same endpoint obtained with the same toxicity test using a laboratory dilution water. Therefore, just as aquatic life criteria can be expressed in terms of either the total recoverable measurement or the dissolved measurement, so can WERs. A pair of side-by-side toxicity tests can produce both a total recoverable WER and a dissolved WER if the metal in the test solutions in both of the tests is measured using both methods. A total recoverable WER is obtained by dividing endpoints that were calculated on the basis of total recoverable metal, whereas a dissolved WER is obtained by dividing endpoints that were calculated on the basis of dissolved metal. Because of the way they are determined, a total recoverable WER is used to calculate a total recoverable site-specific criterion from a national, state, or recalculated aquatic life criterion that is expressed using the total recoverable measurement, whereas a dissolved WER is used to calculate a dissolved site-specific criterion from a national, state, or recalculated criterion that is expressed in terms of the dissolved measurement.

In terms of the classification scheme given in Figure D1, the basic relationship between a total recoverable national water quality criterion and a total recoverable WER is:

- A total recoverable criterion treats all the toxic and nontoxic metal in the site water as if its average toxicity were the same as the average toxicity of all the toxic and nontoxic metal in the toxicity tests in laboratory dilution water on which the criterion is based.
- A total recoverable WER is a measurement of the actual ratio of the average toxicities of the total recoverable metal and replaces the assumption that the ratio is 1.

Similarly, the basic relationship between a dissolved national criterion and a dissolved WER is:

- A dissolved criterion treats all the toxic and nontoxic dissolved metal in the site water as if its average toxicity were the same as the average toxicity of all the toxic and nontoxic dissolved metal in the toxicity tests in laboratory dilution water on which the criterion is based.
- A dissolved WER is a measurement of the actual ratio of the average toxicities of the dissolved metal and replaces the assumption that the ratio is 1.

In both cases, use of a criterion without a WER involves measurement of toxicity in laboratory dilution water but only prediction of toxicity in site water, whereas use of a criterion with a WER involves measurement of toxicity in both laboratory dilution water and site water.

When WERs are used to derive site-specific criteria, the total recoverable and dissolved approaches are inherently consistent. They are consistent because the toxic effects caused by the metal in the toxicity tests do not depend on what chemical measurements are performed; the same number of organisms are killed in the acute lethality tests regardless of what, if any, measurements of the concentration of the metal are made. The only difference is the chemical measurement to which the toxicity is referenced. Dissolved WERs can be derived from the same pairs of toxicity tests from which total recoverable WERs are derived, if the metal in the tests is measured using both the total recoverable and dissolved measurements. Both approaches start at the same place (i.e., the amount of toxicity observed in laboratory dilution water) and end at the same place (i.e., the amount of toxicity observed in site water). The combination of a total recoverable criterion and WER accomplish the same thing as the combination of a dissolved criterion and WER. By extension, whenever a criterion and a WER based on the same measurement of the metal are used together, they will end up at the same place. Because use of a total recoverable criterion with a total recoverable WER ends up at exactly the same place as use of a dissolved criterion with a dissolved WER, whenever one WER is determined, both should be determined to allow (a) a check on the analytical chemistry, (b) use of the inherent internal consistency to check that the data are used correctly, and (c) the option of using either approach in the derivation of permit limits.

An examination of how the two approaches (the total recoverable approach and the dissolved approach) address the four relevant forms of metal (toxic and nontoxic particulate metal and toxic and nontoxic dissolved metal) in laboratory dilution water and in site water further explains why the two approaches are inherently consistent. Here, only the way in which the two approaches address each of the four forms of metal in site water will be considered:

a. Toxic dissolved metal:

This form contributes to the toxicity of the site water and is measured by both chemical measurements. If this is the only form of metal present, the two WERs will be the same.

b. Nontoxic dissolved metal:

This form does not contribute to the toxicity of the site water, but it is measured by both chemical measurements. If this is the only form of metal present, the two WERs will be the same. (Nontoxic dissolved metal can be the only form present, however, only if all of the nontoxic dissolved metal present is refractory. If any labile nontoxic dissolved metal is present, equilibrium will require that some toxic dissolved metal also be present.)

c. Toxic particulate metal:

This form contributes to the toxicological measurement in both approaches; it is measured by the total recoverable measurement, but not by the dissolved measurement. Even though it is not measured by the dissolved measurement, its presence is accounted for in the dissolved approach because it increases the toxicity of the site water and thereby decreases the dissolved WER. It is accounted for because it makes the dissolved metal appear to be more toxic than it is. Most toxic particulate metal is probably not toxic when it is particulate; it becomes toxic when it is dissolved at the gill surface or in the digestive system; in the surface water, however, it is measured as particulate metal.

d. Nontoxic particulate metal:

This form does not contribute to the toxicity of the site water; it is measured by the total recoverable measurement, but not by the dissolved measurement. Because it is measured by the total recoverable measurement, but not by the dissolved measurement, it causes the total recoverable WER to be higher than the dissolved WER.

In addition to dealing with the four forms of metal similarly, the WERs used in the two approaches comparably take synergism, antagonism, and additivity into account. Synergism and additivity in the site water increase its toxicity and therefore decrease the WER; in contrast, antagonism in the site water decreases toxicity and increases the WER.

Each of the four forms of metal is appropriately taken into account because use of the WERs makes the two approaches internally consistent. In addition, although experimental variation will cause the measured WERs to deviate from the actual WERs, the measured WERs will be internally consistent with the data from which they were generated. If the percent dissolved is the same at the test endpoint in the two waters, the two WERs will be the same. If the percent of the total recoverable metal that is dissolved in laboratory dilution water is less than 100 percent, changing from the total recoverable measurement to the dissolved measurement will lower the criterion but it will

comparably lower the denominator in the WER, thus increasing the WER. If the percent of the total recoverable metal that is dissolved in the site water is less than 100 percent, changing from the total recoverable measurement to the dissolved measurement will lower the concentration in the site water that is to be compared with the criterion, but it also lowers the numerator in the WER, thus lowering the WER. Thus when WERs are used to adjust criteria, the total recoverable approach and the dissolved approach result in the same interpretations of concentrations in the site water (see Figure D3) and in the same maximum acceptable concentrations in effluents (see Figure D4).

Thus, if WERs are based on toxicity tests whose endpoints equal the CMC or CCC and if both approaches are used correctly, the two measurements will produce the same results because each WER is based on measurements on the site water and then the WER is used to calculate the site-specific criterion that applies to the site water when the same chemical measurement is used to express the site-specific criterion. The equivalency of the two approaches applies if they are based on the same sample of site water. When they are applied to multiple samples, the approaches can differ depending on how the results from replicate samples are used:

- a. If an appropriate averaging process is used, the two will be equivalent.
- b. If the lowest value is used, the two approaches will probably be equivalent only if the lowest dissolved WER and the lowest total recoverable WER were obtained using the same sample of site water.

There are several advantages to using a dissolved criterion even when a dissolved WER is not used. In some situations use of a dissolved criterion to interpret results of measurements of the concentration of dissolved metal in site water might demonstrate that there is no need to determine either a total recoverable WER or a dissolved WER. This would occur when so much of the total recoverable metal was nontoxic particulate metal that even though the total recoverable criterion was exceeded, the corresponding dissolved criterion was not exceeded. The particulate metal might come from an effluent, a resuspension event, or runoff that washed particulates into the body of water. In such a situation the total recoverable WER would also show that the site-specific criterion was not exceeded, but there would be no need to determine a WER if the criterion were expressed on the basis of the dissolved measurement. If the variation over time in the concentration of particulate metal is much greater than the variation in the concentration of dissolved metal, both the total recoverable concentration and the total recoverable WER are likely to vary so much over time that a dissolved criterion would be much more useful than a total recoverable criterion.

Use of a dissolved criterion without a dissolved WER has three disadvantages, however:

1. Nontoxic dissolved metal in the site water is treated as if it is toxic.
2. Any toxicity due to particulate metal in the site water is ignored.
3. Synergism, antagonism, and additivity in the site water are not taken into account.

Use of a dissolved criterion with a dissolved WER overcomes all three problems. For example, if (a) the total recoverable concentration greatly exceeds the total recoverable criterion, (b) the dissolved concentration is below the dissolved criterion, and (c) there is concern about the possibility of toxicity of particulate metal, the determination of a dissolved WER would demonstrate whether toxicity due to particulate metal is measurable.

Similarly, use of a total recoverable criterion without a total recoverable WER has three comparable disadvantages:

1. Nontoxic dissolved metal in site water is treated as if it is toxic.
2. Nontoxic particulate metal in site water is treated as if it is toxic.
3. Synergism, antagonism, and additivity in site water are not taken into account.

Use of a total recoverable criterion with a total recoverable WER overcomes all three problems. For example, determination of a total recoverable WER would prevent nontoxic particulate metal (as well as nontoxic dissolved metal) in the site water from being treated as if it is toxic.

Relationships between WERs and the Forms of Metals

Probably the best way to understand what WERs can and cannot do is to understand the relationships between WERs and the forms of metals. A WER is calculated by dividing the concentration of a metal that corresponds to a toxicity endpoint in a site water by the concentration of the same metal that corresponds to the same toxicity endpoint in a laboratory dilution water. Therefore, using the classification scheme given in Figure D1:

$$WER = \frac{R_s + N_s + T_s + \Delta N_s + \Delta T_s}{R_L + N_L + T_L + \Delta N_L + \Delta T_L}$$

The subscripts "S" and "L" denote site water and laboratory dilution water, respectively, and:

R = the concentration of Refractory metal in a water. (By definition, all refractory metal is nontoxic metal.)

N = the concentration of Nontoxic labile metal in a water.

T = the concentration of Toxic labile metal in a water.

ΔN = the concentration of metal added during a WER determination that is Nontoxic labile metal after it is added.

ΔT = the concentration of metal added during a WER determination that is Toxic labile metal after it is added.

For a total recoverable WER, each of these five concentrations includes both particulate and dissolved metal, if both are present; for a dissolved WER only dissolved metal is included.

Because the two side-by-side tests use the same endpoint and are conducted under identical conditions with comparable test organisms, $T_S + \Delta T_S = T_L + \Delta T_L$ when the toxic species of the metal are equally toxic in the two waters. If a difference in water quality causes one or more of the toxic species of the metal to be more toxic in one water than the other, or causes a shift in the ratios of various toxic species, we can define

$$H = \frac{T_S + \Delta T_S}{T_L + \Delta T_L}.$$

Thus H is a multiplier that accounts for a proportional increase or decrease in the toxicity of the toxic forms in site water as compared to their toxicities in laboratory dilution water. Therefore, the general WER equation is:

$$WER = \frac{R_S + N_S + \Delta N_S + H(T_L + \Delta T_L)}{R_L + N_L + \Delta N_L + (T_L + \Delta T_L)}.$$

Several things are obvious from this equation:

1. A WER should not be thought of as a simple ratio such as H . H is the ratio of the toxicities of the toxic species of the metal, whereas the WER is the ratio of the sum of the toxic and the nontoxic species of the metal. Only under a very specific set of conditions will $WER = H$. If these conditions are satisfied and if, in addition, $H = 1$, then $WER = 1$. Although it might seem that all of these conditions will rarely be satisfied, it is not all that rare to find that an experimentally determined WER is close to 1.
2. When the concentration of metal in laboratory dilution water is negligible, $R_L = N_L = T_L = 0$ and

$$WER = \frac{R_S + N_S + \Delta N_S + H(\Delta T_L)}{\Delta N_L + \Delta T_L}.$$

Even though laboratory dilution water is low in TOC and TSS, when metals are added to laboratory dilution water in toxicity tests, ions such as hydroxide, carbonate, and chloride react with some metals to form some particulate species and some dissolved species, both of which might be toxic or nontoxic. The metal species that are nontoxic contribute to ΔN_L , whereas those that are toxic contribute to ΔT_L . Hydroxide, carbonate, chloride, TOC, and TSS can increase ΔN_S . Anything that causes ΔN_S to differ from ΔN_L will cause the WER to differ from 1.

3. Refractory metal and nontoxic labile metal in the site water above that in the laboratory dilution water will increase the WER. Therefore, if the WER is determined in downstream water, rather than in upstream water, the WER will be increased by refractory metal and nontoxic labile metal in the effluent.

Thus there are three major reasons why WERs might be larger or smaller than 1:

- a. The toxic species of the metal might be more toxic in one water than in the other, i.e., $H \neq 1$.
- b. ΔN might be higher in one water than in the other.
- c. R and/or N might be higher in one water than in the other.

The last reason might have great practical importance in some situations. When a WER is determined in downstream water, if most of the metal in the effluent is nontoxic, the WER and the endpoint in site water will correlate with the concentration of metal in the site water. In addition, they will depend on the concentration of metal in the effluent and the concentration of effluent in the site water. This correlation will be best for refractory metal because its toxicity cannot be affected by water quality characteristics; even if the effluent and upstream water are quite different so that the water quality characteristics of the site water depend on the percent effluent, the toxicity of the refractory metal will remain constant at zero and the portion of the WER that is due to refractory metal will be additive.

The Dependence of WERs on the Sensitivity of Toxicity Tests

It would be desirable if the magnitude of the WER for a site water were independent of the toxicity test used in the determination of the WER, so that any convenient toxicity test could be used. It can be seen from the general WER equation that the WER will be independent of the toxicity test only if:

$$WER = \frac{H(T_L + \Delta T_L)}{(T_L + \Delta T_L)} = H,$$

which would require that $R_S = N_S = \Delta N_S = R_L = N_L = \Delta N_L = 0$. (It would be easy to assume that $T_L = 0$, but it can be misleading in some situations to make more simplifications than are necessary.)

This is the simplistic concept of a WER that would be advantageous if it were true, but which is not likely to be true very often. Any situation in which one or more of the terms is greater than zero can cause the WER to depend on the sensitivity of the toxicity test, although the difference in the WERs might be small.

Two situations that might be common can illustrate how the WER can depend on the sensitivity of the toxicity test. For these illustrations, there is no advantage to assuming that $H = 1$, so H will be retained for generality.

1. The simplest situation is when $R_s > 0$, i.e., when a substantial concentration of refractory metal occurs in the site water. If, for simplification, it is assumed that $N_s = \Delta N_s = R_L = N_L = \Delta N_L = 0$, then:

$$WER = \frac{R_s + H(T_L + \Delta T_L)}{(T_L + \Delta T_L)} = \frac{R_s}{(T_L + \Delta T_L)} + H.$$

The quantity $T_L + \Delta T_L$ obviously changes as the sensitivity of the toxicity test changes. When $R_s = 0$, then $WER = H$ and the WER is independent of the sensitivity of the toxicity test. When $R_s > 0$, then the WER will decrease as the sensitivity of the test decreases because $T_L + \Delta T_L$ will increase.

2. More complicated situations occur when $(N_s + \Delta N_s) > 0$. If, for simplification, it is assumed that $R_s = R_L = N_L = \Delta N_L = 0$, then:

$$WER = \frac{(N_s + \Delta N_s) + H(T_L + \Delta T_L)}{(T_L + \Delta T_L)} = \frac{(N_s + \Delta N_s)}{(T_L + \Delta T_L)} + H.$$

- a. If $(N_s + \Delta N_s) > 0$ because the site water contains a substantial concentration of a complexing agent that has an affinity for the metal and if complexation converts toxic metal into nontoxic metal, the complexation reaction will control the toxicity of the solution (Allen 1993). A complexation curve can be graphed in several ways, but the S-shaped curve presented in Figure D5 is most convenient here. The vertical axis is "% uncomplexed", which is assumed to correlate with "% toxic". The "% complexed" is then the "% nontoxic". The ratio of nontoxic metal to toxic metal is:

$$\frac{\% \text{nontoxic}}{\% \text{toxic}} = \frac{\% \text{complexed}}{\% \text{uncomplexed}} = v.$$

For the complexed nontoxic metal:

$$v = \frac{\text{concentration of nontoxic metal}}{\text{concentration of toxic metal}}.$$

In the site water, the concentration of complexed nontoxic metal is $(N_s + \Delta N_s)$ and the concentration of toxic metal is $(T_s + \Delta T_s)$, so that:

$$V_s = \frac{(N_s + \Delta N_s)}{(T_s + \Delta T_s)} = \frac{(N_s + \Delta N_s)}{H(T_L + \Delta T_L)} .$$

and

$$WER = \frac{V_s H(T_L + \Delta T_L) + H(T_L + \Delta T_L)}{(T_L + \Delta T_L)} = V_s H + H = H(V_s + 1) .$$

If the WER is determined using a sensitive toxicity test so that the % uncomplexed (i.e., the % toxic) is 10 %, then $V_s = (90 \%)/(10 \%) = 9$, whereas if a less sensitive test is used so that the % uncomplexed is 50 %, then $V_s = (50 \%)/(50 \%) = 1$. Therefore, if a portion of the WER is due to a complexing agent in the site water, the magnitude of the WER can decrease as the sensitivity of the toxicity test decreases because the % uncomplexed will decrease. In these situations, the largest WER will be obtained with the most sensitive toxicity test; progressively smaller WERs will be obtained with less sensitive toxicity tests. The magnitude of a WER will depend not only on the sensitivity of the toxicity test but also on the concentration of the complexing agent and on its binding constant (complexation constant, stability constant). In addition, the binding constants of most complexing agents depend on pH.

If the laboratory dilution water contains a low concentration of a complexing agent,

$$V_L = \frac{N_L + \Delta N_L}{T_L + \Delta T_L}$$

and

$$WER = \frac{V_s H(T_L + \Delta T_L) + H(T_L + \Delta T_L)}{V_L(T_L + \Delta T_L) + (T_L + \Delta T_L)} = \frac{V_s H + H}{V_L + 1} = \frac{H(V_s + 1)}{V_L + 1} .$$

The binding constant of the complexing agent in the laboratory dilution water is probably different from that of the complexing agent in the site water. Although changing from a more sensitive test to a less sensitive test will decrease both V_s and V_L , the amount of effect is not likely to be proportional.

If the change from a more sensitive test to a less sensitive test were to decrease V_L proportionately more than V_s , the change could result in a larger WER, rather

than a smaller WER, as resulted in the case above when it was assumed that the laboratory dilution water did not contain any complexing agent. This is probably most likely to occur if $H \approx 1$ and if $V_s < V_L$, which would mean that $WER < 1$. Although this is likely to be a rare situation, it does demonstrate again the importance of determining WERs using toxicity tests that have endpoints in laboratory dilution water that are close to the CMC or CCC to which the WER is to be applied.

- b. If $(N_s + \Delta N_s) > 0$ because the site water contains a substantial concentration of an ion that will precipitate the metal of concern and if precipitation converts toxic metal into nontoxic metal, the precipitation reaction will control the toxicity of the solution. The "precipitation curve" given in Figure D6 is analogous to the "complexation curve" given in Figure D5; in the precipitation curve, the vertical axis is "% dissolved", which is assumed to correlate with "% toxic". If the endpoint for a toxicity test is below the solubility limit of the precipitate, $(N_s + \Delta N_s) = 0$, whereas if the endpoint for a toxicity test is above the solubility limit, $(N_s + \Delta N_s) > 0$. If WERs are determined with a series of toxicity tests that have increasing endpoints that are above the solubility limit, the WER will reach a maximum value and then decrease. The magnitude of the WER will depend not only on the sensitivity of the toxicity test but also on the concentration of the precipitating agent, the solubility limit, and the solubility of the precipitate.

Thus, depending on the composition of the site water, a WER obtained with an insensitive test might be larger, smaller, or similar to a WER obtained with a sensitive test. Because of the range of possibilities that exist, the best toxicity test to use in the experimental determination of a WER is one whose endpoint in laboratory dilution water is close to the CMC or CCC that is to be adjusted. This is the rationale that was used in the selection of the toxicity tests that are suggested in Appendix I.

The available data indicate that a less sensitive toxicity test usually gives a smaller WER than a more sensitive test (Hansen 1993a). Thus, use of toxicity tests whose endpoints are higher than the CMC or CCC probably will not result in underprotection; in contrast, use of tests whose endpoints are substantially below the CMC or CCC might result in underprotection.

The factors that cause R_s and $(N_s + \Delta N_s)$ to be greater than zero are all external to the test organisms; they are chemical effects that affect the metal in the water. The magnitude of the WER is therefore expected to depend on the toxicity test used only in regard to the sensitivity of the test. If the endpoints for two

different tests occur at the same concentration of the metal, the magnitude of the WERs obtained with the two tests should be the same; they should not depend on (a) the duration of the test, (b) whether the endpoint is based on a lethal or sublethal effect, or (c) whether the species is a vertebrate or an invertebrate.

Another interesting consequence of the chemistry of complexation is that the % uncomplexed will increase if the solution is diluted (Allen and Hansen 1993). The concentration of total metal will decrease with dilution but the % uncomplexed will increase. The increase will not offset the decrease and so the concentration of uncomplexed metal will decrease. Thus the portion of a WER that is due to complexation will not be strictly additive (see Appendix G), but the amount of nonadditivity might be difficult to detect in toxicity studies of additivity. A similar effect of dilution will occur for precipitation.

The illustrations presented above were simplified to make it easier to understand the kinds of effects that can occur. The illustrations are qualitatively valid and demonstrate the direction of the effects, but real-world situations will probably be so much more complicated that the various effects cannot be dealt with separately.

Other Properties of WERs

1. Because of the variety of factors that can affect WERs, no rationale exists at present for extrapolating WERs from one metal to another, from one effluent to another, or from one surface water to another. Thus WERs should be individually determined for each metal at each site.
2. The most important information that the determination of a WER provides is whether simulated and/or actual downstream water adversely affects test organisms that are sensitive to the metal. A WER cannot indicate how much metal needs to be removed from or how much metal can be added to an effluent.
 - a. If the site water already contains sufficient metal that it is toxic to the test organisms, a WER cannot be determined with a sensitive test and so an insensitive test will have to be used. Even if a WER could be determined with a sensitive test, the WER cannot indicate how much metal has to be removed. For example, if a WER indicated that there was 20 percent too much metal in an effluent, a 30 percent reduction by the discharger would not reduce toxicity if only nontoxic metal was removed. The next WER determination would show that the effluent still contained too much metal. Removing metal is useful only if the metal removed is toxic metal. Reducing the total recoverable concentration does not necessarily reduce toxicity.

- b. If the simulated or actual downstream water is not toxic, a WER can be determined and used to calculate how much additional metal the effluent could contain and still be acceptable. Because an unlimited amount of refractory metal can be added to the effluent without affecting the organisms, what the WER actually determines is how much additional toxic metal can be added to the effluent.
3. The effluent component of nearly all WERs is likely to be due mostly to either (a) a reduction in toxicity of the metal by TSS or TOC, or (b) the presence of refractory metal. For both of these, if the percentage of effluent in the downstream water decreases, the magnitude of the WER will usually decrease. If the water quality characteristics of the effluent and the upstream water are quite different, it is possible that the interaction will not be additive; this can affect the portion of the WER that is due to reduced toxicity caused by sorption and/or binding, but it cannot affect the portion of the WER that is due to refractory metal.
4. Test organisms are fed during some toxicity tests, but not during others; it is not clear whether a WER determined in a fed test will differ from a WER determined in an unfed test. Whether there is a difference is likely to depend on the metal, the type and amount of food, and whether a total recoverable or dissolved WER is determined. This can be evaluated by determining two WERs using a test in which the organisms usually are not fed - one WER with no food added to the tests and one with food added to the tests. Any effect of food is probably due to an increase in TOC and/or TSS. If food increases the concentration of nontoxic metal in both the laboratory dilution water and the site water, the food will probably decrease the WER. Because complexes of metals are usually soluble, complexation is likely to lower both total recoverable and dissolved WERs; sorption to solids will probably reduce only total recoverable WERs. The food might also affect the acute-chronic ratio. Any feeding during a test should be limited to the minimum necessary.

Ranges of Actual Measured WERs

The acceptable WERs found by Brungs et al. (1992) were total recoverable WERs that were determined in relatively clean fresh water. These WERs ranged from about 1 to 15 for both copper and cadmium, whereas they ranged from about 0.7 to 3 for zinc. The few WERs that were available for chromium, lead, and nickel ranged from about 1 to 6. Both the total recoverable and dissolved WERs for copper in New York harbor range from about 0.4 to 4 with most of the WERs being between 1 and 2 (Hansen 1993b).

Figure D2: An Example of the Empirical Extrapolation Process

Assume the following hypothetical effluent and upstream water:

Effluent:

T_E : 100 ug/L
 D_E : 10 ug/L (10 % dissolved)
 Q_E : 24 cfs

Upstream water:

T_U : 40 ug/L
 D_U : 38 ug/L (95 % dissolved)
 Q_U : 48 cfs

Downstream water:

T_D : 60 ug/L
 D_D : 36 ug/L (60 % dissolved)
 Q_D : 72 cfs

where:

T = concentration of total recoverable metal.

D = concentration of dissolved metal.

Q = flow.

The subscripts E, U, and D signify effluent, upstream water, and downstream water, respectively.

By conservation of flow: $Q_D = Q_E + Q_U$.

By conservation of total recoverable metal: $T_D Q_D = T_E Q_E + T_U Q_U$.

If P = the percent of the total recoverable metal in the effluent that becomes dissolved in the downstream water,

$$P = \frac{100 (D_D Q_D - D_U Q_U)}{T_E Q_E}$$

For the data given above, the percent of the total recoverable metal in the effluent that becomes dissolved in the downstream water is:

$$P = \frac{100 [(36 \text{ ug/L}) (72 \text{ cfs}) - (38 \text{ ug/L}) (48 \text{ cfs})]}{(100 \text{ ug/L}) (24 \text{ cfs})} = 32 \%$$

which is greater than the 10 % dissolved in the effluent and less than the 60 % dissolved in the downstream water.

Figure D3: The Internal Consistency of the Two Approaches

The internal consistency of the total recoverable and dissolved approaches can be illustrated by considering the use of WERs to interpret the total recoverable and dissolved concentrations of a metal in a site water. For this hypothetical example, it will be assumed that the national CCCs for the metal are:

200 ug/L as total recoverable metal.

160 ug/L as dissolved metal.

It will also be assumed that the concentrations of the metal in the site water are:

300 ug/L as total recoverable metal.

120 ug/L as dissolved metal.

The total recoverable concentration in the site water exceeds the national CCC, but the dissolved concentration does not.

The following results might be obtained if WERs are determined:

In Laboratory Dilution Water

Total recoverable LC50 = 400 ug/L.

% of the total recoverable metal that is dissolved = 80.

(This is based on the ratio of the national CCCs,
which were determined in laboratory dilution water.)

Dissolved LC50 = 320 ug/L.

In Site Water

Total recoverable LC50 = 620 ug/L.

% of the total recoverable metal that is dissolved = 40.

(This is based on the data given above for site water).

Dissolved LC50 = 248 ug/L.

WERs

Total recoverable WER = $(620 \text{ ug/L}) / (400 \text{ ug/L}) = 1.55$

Dissolved WER = $(248 \text{ ug/L}) / (320 \text{ ug/L}) = 0.775$

Checking the Calculations

$$\frac{\text{Total recoverable WER}}{\text{Dissolved WER}} = \frac{1.55}{0.775} = \frac{\text{lab water \% dissolved}}{\text{site water \% dissolved}} = \frac{80}{40} = 2$$

Site-specific CCCs (ssCCCs)

Total recoverable ssCCC = $(200 \text{ ug/L}) (1.55) = 310 \text{ ug/L}$.

Dissolved ssCCC = $(160 \text{ ug/L}) (0.775) = 124 \text{ ug/L}$.

Both concentrations in site water are below the respective ssCCCs.

In contrast, the following results might have been obtained when the WERS were determined:

In Laboratory Dilution Water

Total recoverable LC50 = 400 ug/L.

% of the total recoverable metal that is dissolved = 80.

Dissolved LC50 = 320 ug/L.

In Site Water

Total recoverable LC50 = 580 ug/L.

% of the total recoverable metal that is dissolved = 40.

Dissolved LC50 = 232 ug/L.

WERS

Total recoverable WER = (580 ug/L)/(400 ug/L) = 1.45

Dissolved WER = (232 ug/L)/(320 ug/L) = 0.725

Checking the Calculations

$$\frac{\text{Total recoverable WER}}{\text{Dissolved WER}} = \frac{1.45}{0.725} = \frac{\text{lab water \% dissolved}}{\text{site water \% dissolved}} = \frac{80}{40} = 2$$

Site-specific CCCs (ssCCCs)

Total recoverable ssCCC = (200 ug/L)(1.45) = 290 ug/L.

Dissolved ssCCC = (160 ug/L)(0.725) = 116 ug/L.

In this case, both concentrations in site water are above the respective ssCCCs.

In each case, both approaches resulted in the same conclusion concerning whether the concentration in site water exceeds the site-specific criterion.

The two key assumptions are:

1. The ratio of total recoverable metal to dissolved metal in laboratory dilution water when the WERS are determined equals the ratio of the national CCCs.
2. The ratio of total recoverable metal to dissolved metal in site water when the WERS are determined equals the ratio of the concentrations reported in the site water.

Differences in the ratios that are outside the range of experimental variation will cause problems for the derivation of site-specific criteria and, therefore, with the internal consistency of the two approaches.

Figure D4: The Application of the Two Approaches

Hypothetical upstream water and effluent will be used to demonstrate the equivalence of the total recoverable and dissolved approaches. The upstream water and the effluent will be assumed to have specific properties in order to allow calculation of the properties of the downstream water, which will be assumed to be a 1:1 mixture of the upstream water and effluent. It will also be assumed that the ratios of the forms of the metal in the upstream water and in the effluent do not change when the total recoverable concentration changes.

Upstream water (Flow = 3 cfs)

Total recoverable:	400 ug/L	
Refractory particulate:	200 ug/L	
Toxic dissolved:	200 ug/L	(50 % dissolved)

Effluent (Flow = 3 cfs)

Total recoverable:	440 ug/L	
Refractory particulate:	396 ug/L	
Labile nontoxic particulate:	44 ug/L	
Toxic dissolved:	0 ug/L	(0 % dissolved)

(The labile nontoxic particulate, which is 10 % of the total recoverable in the effluent, becomes toxic dissolved in the downstream water.)

Downstream water (Flow = 6 cfs)

Total recoverable:	420 ug/L	
Refractory particulate:	298 ug/L	
Toxic dissolved:	122 ug/L	(29 % dissolved)

The values for the downstream water are calculated from the values for the upstream water and the effluent:

Total recoverable:	$[3(400) + 3(440)]/6$	$= 420 \text{ ug/L}$
Dissolved:	$[3(200) + 3(44+0)]/6$	$= 122 \text{ ug/L}$
Refractory particulate:	$[3(200) + 3(396)]/6$	$= 298 \text{ ug/L}$

Assumed National CCC (nCCC)

Total recoverable	$= 300 \text{ ug/L}$
Dissolved	$= 240 \text{ ug/L}$

Upstream site-specific CCC (ussCCC)

Assume: Dissolved cccWER = 1.2

Dissolved ussCCC = (1.2) (240 ug/L) = 288 ug/L

By calculation: TR ussCCC = (288 ug/L)/(0.5) = 576 ug/L

Total recoverable cccWER = (576 ug/L)/(300 ug/L) = 1.92

	<u>nCCC</u>	<u>cccWER</u>	<u>ussCCC</u>	<u>Conc.</u>
Total recoverable:	300 ug/L	1.92	576 ug/L	400 ug/L
Dissolved:	240 ug/L	1.2	288 ug/L	200 ug/L
% dissolved	80 %	----	50 %	50 %

Neither concentration exceeds its respective ussCCC.

$$\frac{\text{Total recoverable WER}}{\text{Dissolved WER}} = \frac{1.92}{1.2} = \frac{\text{lab water \% dissolved}}{\text{site water \% dissolved}} = \frac{80}{50} = 1.6$$

Downstream site-specific CCC (dssCCC)

Assume: Dissolved cccWER = 1.8

Dissolved dssCCC = (1.8) (240 ug/L) = 432 ug/L

By calculation: TR dssCCC =

{(432 ug/L - [(200 ug/L)/2])/0.1} + {(400 ug/L)/2} = 3520 ug/L

This calculation determines the amount of dissolved metal contributed by the effluent, accounts for the fact that ten percent of the total recoverable metal in the effluent becomes dissolved, and adds the total recoverable metal contributed by the upstream flow.

Total recoverable cccWER = (3520 ug/L)/(300 ug/L) = 11.73

	<u>nCCC</u>	<u>cccWER</u>	<u>dssCCC</u>	<u>Conc.</u>
Total recoverable:	300 ug/L	11.73	3520 ug/L	420 ug/L
Dissolved:	240 ug/L	1.80	432 ug/L	122 ug/L
% dissolved	80 %	----	12.27 %	29 %

Neither concentration exceeds its respective dssCCC.

$$\frac{\text{Total recoverable WER}}{\text{Dissolved WER}} = \frac{11.73}{1.80} = \frac{\text{lab water \% dissolved}}{\text{site water \% dissolved}} = \frac{80}{12.27} = 6.52$$

Calculating the Maximum Acceptable Concentration in the Effluent

Because neither the total recoverable concentration nor the dissolved concentration in the downstream water exceeds its respective site-specific CCC, the concentration of metal in the effluent could be increased. Under the assumption that the ratios of the two forms of the metal in the effluent do not change when the total recoverable concentration changes, the maximum acceptable concentration of total recoverable metal in the effluent can be calculated as follows:

Starting with the total recoverable dssCCC of 3520 ug/L

$$\frac{(6 \text{ cfs}) (3520 \text{ ug/L}) - (3 \text{ cfs}) (400 \text{ ug/L})}{3 \text{ cfs}} = 6640 \text{ ug/L}$$

Starting with the dissolved dssCCC of 432 ug/L

$$\frac{(6 \text{ cfs}) (432 \text{ ug/L}) - (3 \text{ cfs}) (400 \text{ ug/L}) (0.5)}{(3 \text{ cfs}) (0.10)} = 6640 \text{ ug/L}$$

Checking the Calculations

Total recoverable:

$$\frac{(3 \text{ cfs}) (6640 \text{ ug/L}) + (3 \text{ cfs}) (400 \text{ ug/L})}{6 \text{ cfs}} = 3520 \text{ ug/L}$$

Dissolved:

$$\frac{(3 \text{ cfs}) (6640 \text{ ug/L}) (0.10) + (3 \text{ cfs}) (400 \text{ ug/L}) (0.50)}{6 \text{ cfs}} = 432 \text{ ug/L}$$

The value of 0.10 is used because this is the percent of the total recoverable metal in the effluent that becomes dissolved in the downstream water.

The values of 3520 ug/L and 432 ug/L equal the downstream site-specific CCCs derived above.

Another Way to Calculate the Maximum Acceptable Concentration

The maximum acceptable concentration of total recoverable metal in the effluent can also be calculated from the dissolved dssCCC of 432 ug/L using a partition coefficient to convert from the dissolved dssCCC of 432 ug/L to the total recoverable dssCCC of 3520 ug/L:

$$\frac{[6 \text{ cfs}] \left[\frac{432 \text{ ug/L}}{0.1227} - (3 \text{ cfs}) (400 \text{ ug/L}) \right]}{3 \text{ cfs}} = 6640 \text{ ug/L}$$

Note that the value used for the partition coefficient in this calculation is 0.1227 (the one that applies to the downstream water when the total recoverable concentration of metal in the effluent is 6640 ug/L), not 0.29 (the one that applies when the concentration of metal in the effluent is only 420 ug/L). The three ways of calculating the maximum acceptable concentration give the same result if each is used correctly.

Figure D5: A Generalized Complexation Curve

The curve is for a constant concentration of the complexing ligand and an increasing concentration of the metal.

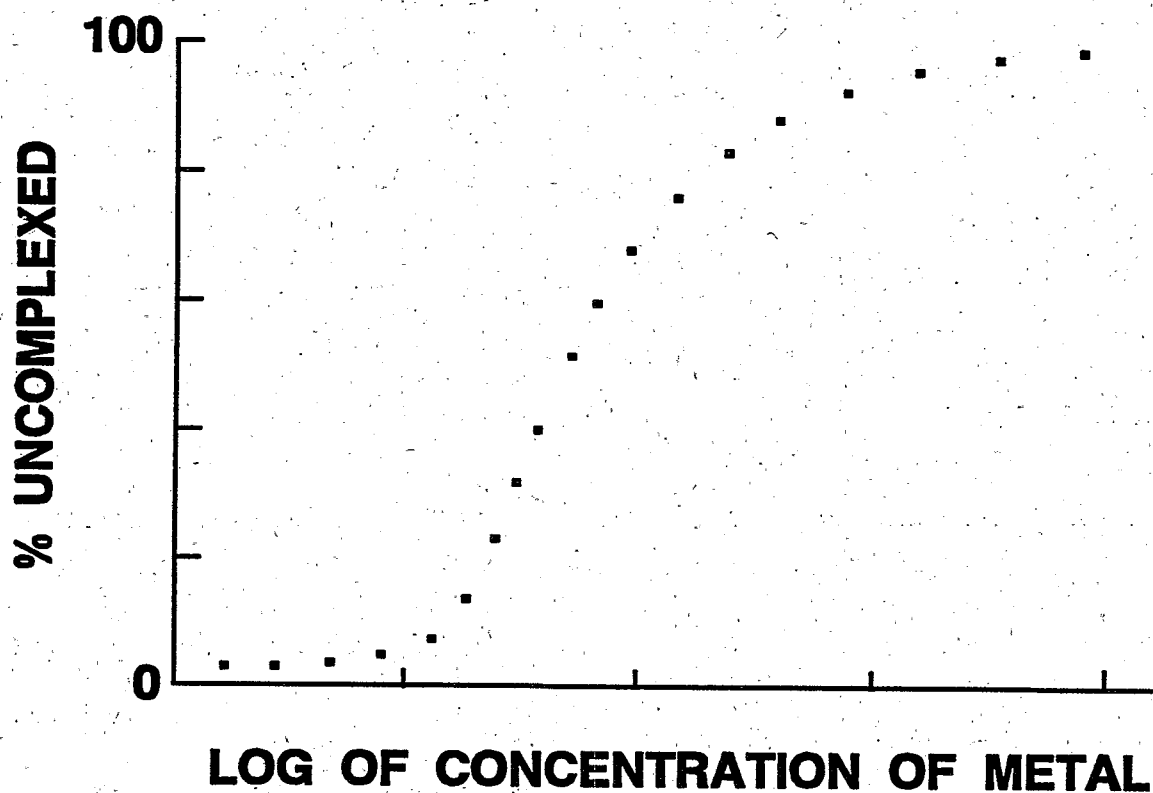
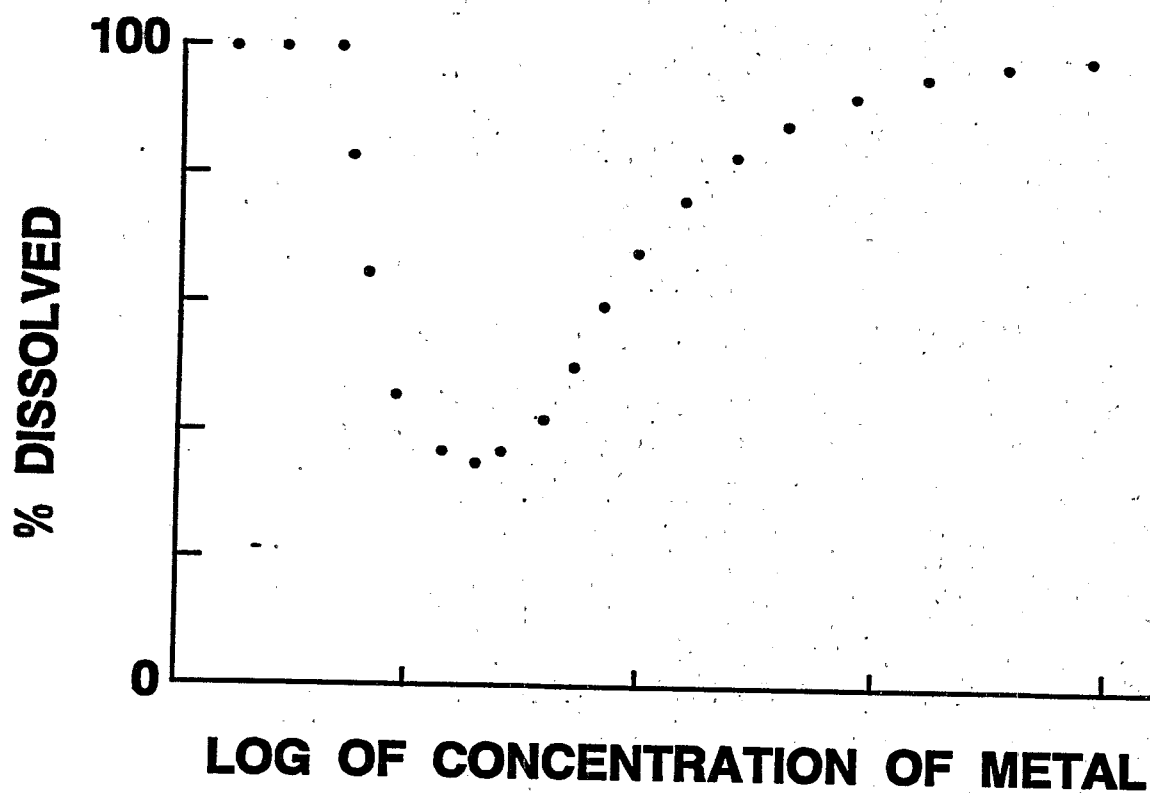


Figure D6: A Generalized Precipitation Curve

The curve is for a constant concentration of the precipitating ligand and an increasing concentration of the metal.



References

- Allen, H.E. 1993. Importance of Metal Speciation to Toxicity. Proceedings of the Water Environment Federation Workshop on Aquatic Life Criteria for Metals. Anaheim, CA. pp. 55-62.
- Allen, H.E., and D.J. Hansen. 1993. The Importance of Trace Metal Speciation to Water Quality Criteria. Paper presented at Society for Environmental Toxicology and Chemistry. Houston, TX. November 15.
- Borgmann, U. 1983. Metal Speciation and Toxicity of Free Metal Ions to Aquatic Biota. IN: Aquatic Toxicology. (J.O. Nriagu, ed.) Wiley, New York, NY.
- Brungs, W.A., T.S. Holderman, and M.T. Southerland. 1992. Synopsis of Water-Effect Ratios for Heavy Metals as Derived for Site-Specific Water Quality Criteria. U.S. EPA Contract 68-CO-0070.
- Chapman, G.A., and J.K. McCrady. 1977. Copper Toxicity: A Question of Form. In: Recent Advances in Fish Toxicology. (R.A. Tubb, ed.) EPA-600/3-77-085 or PB-273 500. National Technical Information Service, Springfield, VA. pp. 132-151.
- Erickson, R. 1993a. Memorandum to C. Stephan. July 14.
- Erickson, R. 1993b. Memorandum to C. Stephan. November 12.
- French, P., and D.T.E. Hunt. 1986. The Effects of Inorganic Complexing upon the Toxicity of Copper to Aquatic Organisms (Principally Fish). IN: Trace Metal Speciation and Toxicity to Aquatic Organisms - A Review. (D.T.E. Hunt, ed.) Report TR 247. Water Research Centre, United Kingdom.
- Hansen, D.J. 1993a. Memorandum to C.E. Stephan. April 29.
- Hansen, D.J. 1993b. Memorandum to C.E. Stephan. October 6.
- Nelson, H., D. Benoit, R. Erickson, V. Mattson, and J. Lindberg. 1986. The Effects of Variable Hardness, pH, Alkalinity, Suspended Clay, and Humics on the Chemical Speciation and Aquatic Toxicity of Copper. PB86-171444. National Technical Information Service, Springfield, VA.
- Wilkinson, K.J., P.M. Bertsch, C.H. Jagoe, and P.G.C. Campbell. 1993. Surface Complexation of Aluminum on Isolated Fish Gill Cells. Environ. Sci. Technol. 27:1132-1138.

Appendix E: U.S. EPA Aquatic Life Criteria Documents for Metals

<u>Metal</u>	<u>EPA Number</u>	<u>NTIS Number</u>
Aluminum	EPA 440/5-86-008	PB88-245998
Antimony	EPA 440/5-80-020	PB81-117319
Arsenic	EPA 440/5-84-033	PB85-227445
Beryllium	EPA 440/5-80-024	PB81-117350
Cadmium	EPA 440/5-84-032	PB85-227031
Chromium	EPA 440/5-84-029	PB85-227478
Copper	EPA 440/5-84-031	PB85-227023
Lead	EPA 440/5-84-027	PB85-227437
Mercury	EPA 440/5-84-026	PB85-227452
Nickel	EPA 440/5-86-004	PB87-105359
Selenium	EPA 440/5-87-006	PB88-142237
Silver	EPA 440/5-80-071	PB81-117822
Thallium	EPA 440/5-80-074	PB81-117848
Zinc	EPA 440/5-87-003	PB87-153581

All are available from:

National Technical Information Service (NTIS)
5285 Port Royal Road
Springfield, VA 22161
TEL: 703-487-4650

Appendix F: Considerations Concerning Multiple-Metal, Multiple-Discharge, and Special Flowing-Water Situations

Multiple-Metal Situations

Both Method 1 and Method 2 work well in multiple-metal situations, although the amount of testing required increases as the number of metals increases. The major problem is the same for both methods: even when addition of two or more metals individually is acceptable, simultaneous addition of the two or more metals, each at its respective maximum acceptable concentration, might be unacceptable for at least two reasons:

1. Additivity or synergism might occur between metals.
2. More than one of the metals might be detoxified by the same complexing agent in the site water. When WERs are determined individually, each metal can utilize all of the complexing capacity; when the metals are added together, however, they cannot simultaneously utilize all of the complexing capacity. Thus a discharger might feel that it is cost-effective to try to justify the lowest site-specific criterion that is acceptable to the discharger rather than trying to justify the highest site-specific criterion that the appropriate regulatory authority might approve.

There are two options for dealing with the possibility of additivity and synergism between metals:

- a. WERs could be developed using a mixture of the metals but it might be necessary to use several primary toxicity tests depending on the specific metals that are of interest. Also, it might not be clear what ratio of the metals should be used in the mixture.
- b. If a WER is determined for each metal individually, one or more additional toxicity tests **must** be conducted at the end to show that the combination of all metals at their proposed new site-specific criteria is acceptable. Acceptability **must** be demonstrated with each toxicity test that was used as a primary toxicity test in the determination of the WERs for the individual metals. Thus if a different primary test was used for each metal, the number of acceptability tests needed would equal the number of metals. It is possible that a toxicity test used as the primary test for one metal might be more sensitive than the CMC (or CCC) for another metal and thus might not be usable in the combination test unless antagonism occurs. When a primary test cannot be used, an acceptable alternative test **must** be used.

The second option is preferred because it is more definitive; it provides data for each metal individually and for the mixture. The first option leaves the possibility that one of the metals is antagonistic towards another so that the toxicity of the mixture would increase if the metal causing the antagonism were not present.

Multiple-Discharge Situations

Because the National Toxics Rule (NTR) incorporated WERs into the aquatic life criteria for some metals, it might be envisioned that more than one criterion could apply to a metal at a site if different investigators obtained different WERs for the same metal at the site. In jurisdictions subject to the NTR, as well as in all other jurisdictions, EPA intends that there should be no more than one criterion for a pollutant at a point in a body of water. Thus whenever a site-specific criterion is to be derived using a WER at a site at which more than one discharger has permit limits for the same metal, it is important that all dischargers work together with the appropriate regulatory authority to develop a workplan that is designed to derive a site-specific criterion that adequately protects the entire site.

Method 2 is ideally suited for taking into account more than one discharger.

Method 1 is straightforward if the dischargers are sufficiently far downstream of each other that the stream can be divided into a separate site for each discharger. Method 1 can also be fairly straightforward if the WERs are additive, but it will be complex if the WERs are not additive. Deciding whether to use a simulated downstream water or an actual downstream water can be difficult in a flowing-water multiple-discharge situation. Use of actual downstream water can be complicated by the existence of multiple mixing zones and plumes and by the possibility of varying discharge schedules; these same problems exist, however, if effluents from two or more discharges are used to prepare simulated downstream water. Dealing with a multiple-discharge situation is much easier if the WERs are additive, and use of simulated downstream water is the best way to determine whether the WERs are additive. Taking into account all effluents will take into account synergism, antagonism, and additivity. If one of the discharges stops or is modified substantially, however, it will usually be necessary to determine a new WER, except possibly if the metal being discharged is refractory. Situations concerning intermittent and batch discharges need to be handled on a case-by-case basis.

Special Flowing-Water Situations

Method 1 is intended to apply not only to ordinary rivers and streams but also to streams that some people might consider "special", such as streams whose design flows are zero and streams that some state and/or federal agencies might refer to as "effluent-dependent", "habitat-creating", "effluent-dominated", etc. (Due to differences between agencies, some streams whose design flows are zero are not considered "effluent-dependent",

etc., and some "effluent-dependent" streams have design flows that are greater than zero.) The application of Method 1 to these kinds of streams has the following implications:

1. If the design flow is zero, at least some WERs ought to be determined in 100% effluent.
2. If thunderstorms, etc., occasionally dilute the effluent substantially, at least one WER should be determined in diluted effluent to assess whether dilution by rainwater might result in underprotection by decreasing the WER faster than it decreases the concentration of the metal. This might occur, for example, if rainfall reduces hardness, alkalinity, and pH substantially. This might not be a concern if the WER demonstrates a substantial margin of safety.
3. If the site-specific criterion is substantially higher than the national criterion, there should be increased concern about the fate of the metal that has reduced or no toxicity. Even if the WER demonstrates a substantial margin of safety (e.g., if the site-specific criterion is three times the national criterion, but the experimentally determined WER is 11), it might be desirable to study the fate of the metal.
4. If the stream merges with another body of water and a site-specific criterion is desired for the merged waters, another WER needs to be determined for the mixture of the waters.
5. Whether WET testing is required is not a WER issue, although WET testing might be a condition for determining and/or using a WER.
6. A concern about what species should be present and/or protected in a stream is a beneficial-use issue, not a WER issue, although resolution of this issue might affect what species should be used if a WER is determined. (If the Recalculation Procedure is used, determining what species should be present and/or protected is obviously important.)
7. Human health and wildlife criteria and other issues might restrict an effluent more than an aquatic life criterion.

Although there are no scientific reasons why "effluent-dependent", etc., streams and streams whose design flows are zero should be subject to different guidance than other streams, a regulatory decision (for example, see 40 CFR 131) might require or allow some or all such streams to be subject to different guidance. For example, it might be decided on the basis of a use attainability analysis that one or more constructed streams do not have to comply with usual aquatic life criteria because it is decided that the water quality in such streams does not need to protect sensitive aquatic species. Such a decision might eliminate any further concern for site-specific aquatic life criteria and/or for WET testing for such streams. The water quality might be unacceptable for other reasons, however.

In addition to its use with rivers and streams, Method 1 is also appropriate for determining cmcWERs that are applicable to near-field effects of discharges into large bodies of fresh or salt water, such as an ocean or a large lake, reservoir, or estuary:

- a. The near-field effects of a pipe that extends far into a large body of fresh or salt water that has a current, such as an ocean, can probably best be treated the same as a single discharge into a flowing stream. For example, if a mixing zone is defined, the concentration of effluent at the edge of the mixing zone might be used to define how to prepare a simulated site water. A dye dispersion study (Kilpatrick 1992) might be useful, but a dilution model (U.S. EPA 1993) is likely to be a more cost-effective way of obtaining information concerning the amount of dilution at the edge of the mixing zone.
- b. The near-field effects of a single discharge that is near a shore of a large body of fresh or salt water can also probably best be treated the same as a single discharge into a flowing stream, especially if there is a definite plume and a defined mixing zone. The potential point of impact of near-field effects will often be an embayment, bayou, or estuary that is a nursery for fish and invertebrates and/or contains commercially important shellfish beds. Because of their importance, these areas should receive special consideration in the determination and use of a WER, taking into account sources of water and discharges, mixing patterns, and currents (and tides in coastal areas). The current and flushing patterns in estuaries can result in increased pollutant concentrations in confined embayments and at the terminal up-gradient portion of the estuary due to poor tidal flushing and exchange. Dye dispersion studies (Kilpatrick 1992) can be used to determine the spatial concentration of the effluent in the receiving water, but dilution models (U.S. EPA 1993) might not be sufficiently accurate to be useful. Dye studies of discharges in near-shore tidal areas are especially complex. Dye injection into the discharge should occur over at least one, and preferably two or three, complete tidal cycles; subsequent dispersion patterns should be monitored in the ambient water on consecutive tidal cycles using an intensive sampling regime over time, location, and depth. Information concerning dispersion and the community at risk can be used to define the appropriate mixing zone(s), which might be used to define how to prepare simulated site water.

References

- Kilpatrick, F.A. 1992. Simulation of Soluble Waste Transport and Buildup in Surface Waters Using Tracers. Open-File Report 92-457. U.S. Geological Survey, Books and Open-File Reports, Box 25425, Federal Center, Denver, CO 80225.
- U.S. EPA. 1993. Dilution Models for Effluent Discharges. Second Edition. EPA/600/R-93/139. National Technical Information Service, Springfield, VA.

Appendix G: Additivity and the Two Components of a WER Determined Using Downstream Water

The Concept of Additivity of WERS

In theory, whenever samples of effluent and upstream water are taken, determination of a WER in 100 % effluent would quantify the effluent WER (eWER) and determination of a WER in 100 % upstream water would quantify the upstream WER (uWER); determination of WERS in known mixtures of the two samples would demonstrate whether the eWER and the uWER are additive. For example, if eWER = 40, uWER = 5, and the two WERS are additive, a mixture of 20 % effluent and 80 % upstream water would give a WER of 12, except possibly for experimental variation, because:

$$\frac{20(eWER) + 80(uWER)}{100} = \frac{20(40) + 80(5)}{100} = \frac{800 + 400}{100} = \frac{1200}{100} = 12 .$$

Strict additivity of an eWER and an uWER will probably be rare because one or both WERS will probably consist of a portion that is additive and a portion that is not. The portions of the eWER and uWER that are due to refractory metal will be strictly additive, because a change in water quality will not make the metal more or less toxic. In contrast, metal that is nontoxic because it is complexed by a complexing agent such as EDTA will not be strictly additive because the % uncomplexed will decrease as the solution is diluted; the amount of change in the % uncomplexed will usually be small and will depend on the concentration and the binding constant of the complexing agent (see Appendix D). Whether the nonrefractory portions of the uWER and eWER are additive will probably also depend on the differences between the water quality characteristics of the effluent and the upstream water, because these will determine the water quality characteristics of the downstream water. If, for example, 85 % of the eWER and 30 % of the uWER are due to refractory metal, the WER obtained in the mixture of 20 % effluent and 80 % upstream water could range from 8 to 12. The WER of 8 would be obtained if the only portions of the eWER and uWER that are additive are those due to refractory metal, because:

$$\frac{20(0.85)(eWER) + 80(0.30)(uWER)}{100} = \frac{20(0.85)(40) + 80(0.30)(5)}{100} = 8 .$$

The WER could be as high as 12 depending on the percentages of the other portions of the WERS that are also additive. Even if the eWER and uWER are not strictly additive, the concept of additivity of WERS can be useful insofar as the eWER and uWER are partially additive, i.e., insofar as a portion of at least one of the WERS is additive. In the example given above, the WER determined using downstream water that consisted of 20 % effluent

and 80 % upstream water would be 12 if the eWER and uWER were strictly additive; the downstream WER would be less than 12 if the eWER and uWER were partially additive.

The Importance of Additivity

The major advantage of additivity of WERs can be demonstrated using the effluent and upstream water that were used above. To simplify this illustration, the acute-chronic ratio will be assumed to be large, and the eWER of 40 and the uWER of 5 will be assumed to be cccWERs that will be assumed to be due to refractory metal and will therefore be strictly additive. In addition, the complete-mix downstream water at design-flow conditions will be assumed to be 20 % effluent and 80 % upstream water, so that the downstream WER will be 12 as calculated above for strict additivity.

Because the eWER and the uWER are cccWERs and are strictly additive, this metal will cause neither acute nor chronic toxicity in downstream water if (a) the concentration of metal in the effluent is less than 40 times the CCC and (b) the concentration of metal in the upstream water is less than 5 times the CCC. As the effluent is diluted by mixing with upstream water, both the eWER and the concentration of metal will be diluted simultaneously; proportional dilution of the metal and the eWER will prevent the metal from causing acute or chronic toxicity at any dilution. When the upstream flow equals the design flow, the WER in the plume will decrease from 40 at the end of the pipe to 12 at complete mix as the effluent is diluted by upstream water; because this WER is due to refractory metal, neither fate processes nor changes in water quality characteristics will affect the WER. When stream flow is higher or lower than design flow, the complete-mix WER will be lower or higher, respectively, than 12, but toxicity will not occur because the concentration of metal will also be lower or higher.

If the eWER and the uWER are strictly additive and if the national CCC is 1 mg/L, the following conclusions are valid when the concentration of the metal in 100 % effluent is less than 40 mg/L and the concentration of the metal in 100 % upstream water is less than 5 mg/L:

1. This metal will not cause acute or chronic toxicity in the upstream water, in 100 % effluent, in the plume, or in downstream water.
2. There is no need for an acute or a chronic mixing zone where a lesser degree of protection is provided.
3. If no mixing zone exists, there is no discontinuity at the edge of a mixing zone where the allowed concentration of metal decreases instantaneously.

These results also apply to partial additivity as long as the concentration of metal does not exceed that allowed by the amount

of additivity that exists. It would be more difficult to take into account the portions of the eWER and uWER that are not additive.

The concept of additivity becomes unimportant when the ratios, concentrations of the metals, or WERs are very different. For example, if eWER = 40, uWER = 5, and they are additive, a mixture of 1 % effluent and 99 % upstream water would have a WER of 5.35. Given the reproducibility of toxicity tests and WERs, it would be extremely difficult to distinguish a WER of 5 from a WER of 5.35. In cases of extreme dilution, rather than experimentally determining a WER, it is probably acceptable to use the limiting WER of 5 or to calculate a WER if additivity has been demonstrated.

Traditionally it has been believed that it is environmentally conservative to use a WER determined in upstream water (i.e., the uWER) to derive a site-specific criterion that applies downstream (i.e., that applies to areas that contain effluent). This belief is probably based on the assumption that a larger WER would be obtained in downstream water that contains effluent, but the belief could also be based on the assumption that the uWER is additive. It is possible that in some cases neither assumption is true, which means that using a uWER to derive a downstream site-specific criterion might result in underprotection. It seems likely, however, that WERs determined using downstream water will usually be at least as large as the uWER.

Several kinds of concerns about the use of WERs are actually concerns about additivity:

1. Do WERs need to be determined at higher flows in addition to being determined at design flow?
2. Do WERs need to be determined when two bodies of water mix?
3. Do WERs need to be determined for each additional effluent in a multiple-discharge situation.

In each case, the best use of resources might be to test for additivity of WERs.

Mixing Zones

In the example presented above, there would be no need for a regulatory mixing zone with a reduced level of protection if:

1. The eWER is always 40 and the concentration of the metal in 100 % effluent is always less than 40 mg/L.
2. The uWER is always 5 and the concentration of the metal in 100 % upstream water is always less than 5 mg/L.
3. The WERs are strictly additive.

If, however, the concentration exceeded 40 mg/L in 100 % effluent, but there is some assimilative capacity in the upstream water, a regulatory mixing zone would be needed if the discharge were to be allowed to utilize some or all of the assimilative

capacity. The concept of additivity of WERs can be used to calculate the maximum allowed concentration of the metal in the effluent if the eWER and the uWER are strictly additive.

If the concentration of metal in the upstream water never exceeds 0.8 mg/L, the discharger might want to determine how much above 40 mg/L the concentration could be in 100 % effluent. If, for example, the downstream water at the edge of the chronic mixing zone under design-flow conditions consists of 70 % effluent and 30 % upstream water, the WER that would apply at the edge of the mixing zone would be:

$$\frac{70(eWER) + 30(uWER)}{100} = \frac{70(40) + 30(5)}{100} = \frac{2800 + 150}{100} = 29.5 .$$

Therefore, the maximum concentration allowed at this point would be 29.5 mg/L. If the concentration of the metal in the upstream water was 0.8 mg/L, the maximum concentration allowed in 100 % effluent would be 41.8 mg/L because:

$$\frac{70(41.8 \text{ mg/L}) + 30(0.8 \text{ mg/L})}{100} = \frac{2926 \text{ mg/L} + 24 \text{ mg/L}}{100} = 29.5 \text{ mg/L} .$$

Because the eWER is 40, if the concentration of the metal in 100 % effluent is 41.8 mg/L, there would be chronic toxicity inside the chronic mixing zone. If the concentration in 100 % effluent is greater than 41.8 mg/L, there would be chronic toxicity past the edge of the chronic mixing zone. Thus even if the eWER and the uWER are taken into account and they are assumed to be completely additive, a mixing zone is necessary if the assimilative capacity of the upstream water is used to allow discharge of more metal.

If the complete-mix downstream water consists of 20 % effluent and 80 % upstream water at design flow, the complete-mix WER would be 12 as calculated above. The complete-mix approach to determining and using downstream WERs would allow a maximum concentration of 12 mg/L at the edge of the chronic mixing zone, whereas the alternative approach resulted in a maximum allowed concentration of 29.5 mg/L. The complete-mix approach would allow a maximum concentration of 16.8 mg/L in the effluent because:

$$\frac{70(16.8 \text{ mg/L}) + 30(0.8 \text{ mg/L})}{100} = \frac{1176 \text{ mg/L} + 24 \text{ mg/L}}{100} = 12 \text{ mg/L} .$$

In this example, the complete-mix approach limits the concentration of the metal in the effluent to 16.8 mg/L, even though it is known that as long as the concentration in 100 % effluent is less than 40 mg/L, chronic toxicity will not occur inside or outside the mixing zone. If the WER of 12 is used to derive a site-specific CCC of 12 mg/L that is applied to a site

that starts at the edge of the chronic mixing zone and extends all the way across the stream, there would be overprotection at the edge of the chronic mixing zone (because the maximum allowed concentration is 12 mg/L, but a concentration of 29.5 mg/L will not cause chronic toxicity), whereas there would be underprotection on the other side of the stream (because the maximum allowed concentration is 12 mg/L, but concentrations above 5 mg/L can cause chronic toxicity.)

The Experimental Determination of Additivity

Experimental variation makes it difficult to quantify additivity without determining a large number of WERs, but the advantages of demonstrating additivity might be sufficient to make it worth the effort. It should be possible to decide whether the eWER and uWER are strictly additive based on determination of the eWER in 100 % effluent, determination of the uWER in 100 % upstream water, and determination of WERs in 1:3, 1:1, and 3:1 mixtures of the effluent and upstream water, i.e., determination of WERs in 100, 75, 50, 25, and 0 % effluent. Validating models of partial additivity and/or interactions will probably require determination of more WERs and more sophisticated data analysis (see, for example, Broderius 1991).

In some cases chemical measurements or manipulations might help demonstrate that at least some portion of the eWER and/or the uWER is additive:

1. If the difference between the dissolved WER and the total recoverable WER is explained by the difference between the dissolved and total recoverable concentrations, the difference is probably due to particulate refractory metal.
2. If the WERs in different samples of the effluent correlate with the concentration of metal in the effluent, all, or nearly all, of the metal in the effluent is probably nontoxic.
3. A WER that remains constant as the pH is lowered to 6.5 and raised to 9.0 is probably additive.

The concentration of refractory metal is likely to be low in upstream water except during events that increase TSS and/or TOC; the concentration of refractory metal is more likely to be substantial in effluents. Chemical measurements might help identify the percentages of the eWER and the uWER that are due to refractory metal, but again experimental variation will limit the usefulness of chemical measurements when concentrations are low.

Summary

The distinction between the two components of a WER determined using downstream water has the following implications:

1. The magnitude of a WER determined using downstream water will usually depend on the percent effluent in the sample.

2. Insofar as the eWER and uWER are additive, the magnitude of a downstream WER can be calculated from the eWER, the uWER, and the ratio of effluent and upstream water in the downstream water.
3. The derivation and implementation of site-specific criteria should ensure that each component is applied only where it occurs.
 - a. Underprotection will occur if, for example, any portion of the eWER is applied to an area of a stream where the effluent does not occur.
 - b. Overprotection will occur if, for example, an unnecessarily small portion of the eWER is applied to an area of a stream where the effluent occurs.
4. Even though the concentration of metal might be higher than a criterion in both a regulatory mixing zone and a plume, a reduced level of protection is allowed in a mixing zone, whereas a reduced level of protection is not allowed in the portion of a plume that is not inside a mixing zone.
5. Regulatory mixing zones are necessary if, and only if, a discharger wants to make use of the assimilative capacity of the upstream water.
6. It might be cost-effective to quantify the eWER and uWER, determine the extent of additivity, study variability over time, and then decide how to regulate the metal in the effluent.

Reference

Broderius, S.J. 1991. Modeling the Joint Toxicity of Xenobiotics to Aquatic Organisms: Basic Concepts and Approaches. In: Aquatic Toxicology and Risk Assessment: Fourteenth Volume. (M.A. Mayes and M.G. Barron, eds.) ASTM STP 1124. American Society for Testing and Materials, Philadelphia, PA. pp. 107-127.

Appendix H: Special Considerations Concerning the Determination of WERs with Saltwater Species

1. The test organisms should be compatible with the salinity of the site water, and the salinity of the laboratory dilution water should match that of the site water. Low-salinity stenohaline organisms should not be tested in high-salinity water, whereas high-salinity stenohaline organisms should not be tested in low-salinity water; it is not known, however, whether an incompatibility will affect the WER. If the community to be protected principally consists of euryhaline species, the primary and secondary toxicity tests should use the euryhaline species suggested in Appendix I (or taxonomically related species) whenever possible, although the range of tolerance of the organisms should be checked.
 - a. When Method 1 is used to determine cmcWERS at saltwater sites, the selection of test organisms is complicated by the fact that most effluents are freshwater and they are discharged into salt waters having a wide range of salinities. Some state water quality standards require a permittee to meet an LC50 or other toxicity limit at the end of the pipe using a freshwater species. However, the intent of the site-specific and national water quality criteria program is to protect the communities that are at risk. Therefore, freshwater species should not be used when WERs are determined for saltwater sites unless such freshwater species (or closely related species) are in the community at risk. The addition of a small amount of brine and the use of salt-tolerant freshwater species is inappropriate for the same reason. The addition of a large amount of brine and the use of saltwater species that require high salinity should also be avoided when salinity is likely to affect the toxicity of the metal. Salinities that are acceptable for testing euryhaline species can be produced by dilution of effluent with sea water and/or addition of a commercial sea salt or a brine that is prepared by evaporating site water; small increases in salinity are acceptable because the effluent will be diluted with salt water wherever the communities at risk are exposed in the real world. Only as a last resort should freshwater species that tolerate low levels of salinity and are sensitive to metals, such as Daphnia magna and Hyalella azteca, be used.
 - b. When Method 2 is used to determine cccWERS at saltwater sites:
 - 1) If the site water is low-salinity but all the sensitive test organisms are high-salinity stenohaline organisms, a commercial sea salt or a brine that is prepared by evaporating site water may be added in order to increase the salinity to the minimum level that is acceptable to the test organisms; it should be determined whether the

salt or brine reduces the toxicity of the metal and thereby increases the WER.

- 2) If the site water is high-salinity, selecting test organisms should not be difficult because many of the sensitive test organisms are compatible with high-salinity water.
2. It is especially important to consider the availability of test organisms when saltwater species are to be used, because many of the commonly used saltwater species are not cultured and are only available seasonally.
3. Many standard published methodologies for tests with saltwater species recommend filtration of dilution water, effluent, and/or test solutions through a 37- μ m sieve or screen to remove predators. Site water should be filtered only if predators are observed in the sample of the water because filtration might affect toxicity. Although recommended in some test methodologies, ultraviolet treatment is often not needed and generally should be avoided.
4. If a natural salt water is to be used as the laboratory dilution water, the samples should probably be collected at slack high tide (\pm 2 hours). Unless there is stratification, samples should probably be taken at mid-depth; however, if a water quality characteristic, such as salinity or TSS, is important, the vertical and horizontal definition of the point of sampling might be important. A conductivity meter, salinometer, and/or transmissometer might be useful for determining where and at what depth to collect the laboratory dilution water; any measurement of turbidity will probably correlate with TSS.
5. The salinity of the laboratory dilution water should be within \pm 10 percent or 2 mg/L (whichever is higher) of that of the site water.

Appendix I: Suggested Toxicity Tests for Determining WERs for Metals

Selecting primary and secondary toxicity tests for determining WERs for metals should take into account the following:

1. WERs determined with more sensitive tests are likely to be larger than WERs determined with less sensitive tests (see Appendix D). Criteria are derived to protect sensitive species and so WERs should be derived to be appropriate for sensitive species. The appropriate regulatory authority will probably accept WERs derived with less sensitive tests because such WERs are likely to provide at least as much protection as WERs determined with more sensitive tests.
2. The species used in the primary and secondary tests **must** be in different orders and should include a vertebrate and an invertebrate.
3. The test organism (i.e., species and life stage) should be readily available throughout the testing period.
4. The chances of the test being successful should be high.
5. The relative sensitivities of test organisms vary substantially from metal to metal.
6. The sensitivity of a species to a metal usually depends on both the life stage and kind of test used.
7. Water quality characteristics might affect chronic toxicity differently than they affect acute toxicity (Spehar and Carlson 1984; Chapman, unpublished; Voyer and McGovern 1991).
8. The endpoint of the primary test in laboratory dilution water should be as close as possible (but **must not** be below) the CMC or CCC to which the WER is to be applied; the endpoint of the secondary test should be as close as possible (and should not be below) the CMC or CCC.
9. Designation of tests as acute and chronic has no bearing on whether they may be used to determine a cmcWER or a cccWER. The suggested toxicity tests should be considered, but the actual selection should depend on the specific circumstances that apply to a particular WER determination.

Regardless of whether test solutions are renewed when tests are conducted for other purposes, if the concentrations of dissolved metal and dissolved oxygen remain acceptable when determining WERs, tests whose duration is not longer than 48 hours may be static tests, whereas tests whose duration is longer than 48 hours **must** be renewal tests. If the concentration of dissolved metal and/or the concentration of dissolved oxygen does not remain acceptable, the test solutions **must** be renewed every 24 hours. If one test in a pair of side-by-side tests is a renewal test, both of the tests **must** be renewed on the same schedule.

Appendix H should be read if WERs are to be determined with saltwater species.

Suggested Tests¹ for Determining cmcWERS and cccWERS².
(Concentrations are to be measured in all tests.)

<u>Metal</u>	<u>Water</u> ³	<u>cmcWERS</u> ⁴		<u>cccWERS</u> ⁴	
Aluminum	FW	DA	X	CDC	X
Arsenic (III)	FW	DA	GM	CDC	FMC
	SW	BM	CR	MYC	BM
Cadmium	FW	DA	SL ⁵ or FM	CDC	FMC
	SW	MY	CR	MYC	X
Chrom (III)	FW	GM	SL or DA	FMC	CDC
Chrom (VI)	FW	DA	GM	CDC	GM
	SW	MY	NE	MYC	NEC
Copper	FW	DA	FM or GM	CDC	FM
	SW	BM	AR	BMC	AR
Lead	FW	DA	GM	CDC	X
	SW	BM	MYC	MYC	X
Mercury	FW	DA	GM	Y	Y
	SW	MY	BM	Y	Y
Nickel	FW	DA	FX	CDC	FMC
	SW	MY	BM	MYC	BMC
Selenium	FW	Y	Y	Y	Y
	SW	CR	MYC	MYC	X
Silver	FW	DA	FMC	CDC	FMC
	SW	BM	CR	MYC	BMC
Zinc	FW	DA	FM	CDC	FMC
	SW	BM	MY	MYC	BMC

¹ The description of a test specifies not only the test species and the duration of the test but also the life stage of the species and the adverse effect(s) on which the endpoint is to be based.

² Some tests that are sensitive and are used in criteria documents are not suggested here because the chances of the test organisms being available and the test being successful might be low. Such tests may be used if desired.

- ³ FW = Fresh Water; SW = Salt Water.
- ⁴ Two-letter codes are used for acute tests, whereas codes for chronic tests contain three letters and end in "C". One-letter codes are used for comments.
- ⁵ In acute tests on cadmium with salmonids, substantial numbers of fish usually die after 72 hours. Also, the fish are sensitive to disturbance, and it is sometimes difficult to determine whether a fish is dead or immobilized.

ACUTE TESTS

- AR. A 48-hr EC50 based on mortality and abnormal development from a static test with embryos and larvae of sea urchins of a species in the genus Arbacia (ASTM 1993a) or of the species Strongylocentrotus purpuratus (Chapman 1992).
- BM. A 48-hr EC50 based on mortality and abnormal larval development from a static test with embryos and larvae of a species in one of four genera (Crassostrea, Mulinia, Mytilus, Mercenaria) of bivalve molluscs (ASTM 1993b).
- CR. A 48-hr EC50 (or LC50 if there is no immobilization) from a static test with Acartia or larvae of a saltwater crustacean; if molting does not occur within the first 48 hours, renew at 48 hours and continue the test to 96 hours (ASTM 1993a).
- DA. A 48-hr EC50 (or LC50 if there is no immobilization) from a static test with a species in one of three genera (Ceriodaphnia, Daphnia, Simocephalus) in the family Daphnidae (U.S. EPA 1993a; ASTM 1993a).
- FM. A 48-hr LC50 from a static test at 25°C with fathead minnow (Pimephales promelas) larvae that are 1 to 24 hours old (ASTM 1993a; U.S. EPA 1993a). The embryos **must** be hatched in the laboratory dilution water, except that organisms to be used in the site water may be hatched in the site water. The larvae **must not** be fed before or during the test and at least 90 percent **must** survive in laboratory dilution water for at least six days after hatch.

Note: The following 48-hr LC50s were obtained at a hardness of 50 mg/L with fathead minnow larvae that were 1 to 24 hours old. The metal was measured using the total recoverable procedure (Peltier 1993):

<u>Metal</u>	<u>LC50 (µg/L)</u>
Cadmium	13.87
Copper	6.33
Zinc	100.95

FX. A 96-hr LC50 from a renewal test (renew at 48 hours) at 25°C with fathead minnow (Pimephales promelas) larvae that are 1 to 24 hours old (ASTM 1993a; U.S. EPA 1993a). The embryos **must** be hatched in the laboratory dilution water, except that organisms to be used in the site water may be hatched in the site water. The larvae **must not** be fed before or during the test and at least 90 percent **must** survive in laboratory dilution water for at least six days after hatch.

Note: A 96-hr LC50 of 188.14 µg/L was obtained at a hardness of 50 mg/L in a test on nickel with fathead minnow larvae that were 1 to 24 hours old. The metal was measured using the total recoverable procedure (Peltier 1993). A 96-hr LC50 is used for nickel because substantial mortality occurred after 48 hours in the test on nickel, but not in the tests on cadmium, copper, and zinc.

GM. A 96-hr EC50 (or LC50 if there is no immobilization) from a renewal test (renew at 48 hours) with a species in the genus Gammarus (ASTM 1993a).

MY. A 96-hr EC50 (or LC50 if there is no immobilization) from a renewal test (renew at 48 hours) with a species in one of two genera (Mysidopsis, Holmesimysis [nee Acanthomysis]) in the family Mysidae (U.S. EPA 1993a; ASTM 1993a). Feeding is required during all acute and chronic tests with mysids; for determining WERs, mysids should be fed four hours before the renewal at 48 hours and minimally on the non-renewal days.

NE. A 96-hr LC50 from a renewal test (renew at 48 hours) using juvenile or adult polychaetes in the genus Nereidae (ASTM 1993a).

SL. A 96-hr EC50 (or LC50 if there is no immobilization) from a renewal test (renew at 48 hours) with a species in one of two genera (Oncorhynchus, Salmo) in the family Salmonidae (ASTM 1993a).

CHRONIC TESTS

BMC. A 7-day IC25 from a survival and development renewal test (renew every 48 hours) with a species of bivalve mollusc, such as a species in the genus Mulinia. One such test has been described by Burgess et al. 1992. [Note: When determining WERs, sediment **must not** be in the test chamber.] [Note: This test has not been widely used.]

CDC. A 7-day IC25 based on reduction in survival and/or reproduction in a renewal test with a species in the genus Ceriodaphnia in the family Daphnidae (U.S. EPA 1993b). The

test solutions **must** be renewed every 48 hours. (A 21-day life-cycle test with Daphnia magna is also acceptable.)

FMC. A 7-day IC25 from a survival and growth renewal test. (renew every 48 hours) with larvae (≤ 48 -hr old) of the fathead minnow (Pimephales promelas) (U.S. EPA 1993b). When determining WERs, the fish **must** be fed four hours before each renewal and minimally during the non-renewal days.

MYC. A 7-day IC25 based on reduction in survival, growth, and/or reproduction in a renewal test with a species in one of two genera (Mysidopsis, Holmesimysis [nee Acanthomysis]) in the family Mysidae (U.S. EPA 1993c). Mysids **must** be fed during all acute and chronic tests; when determining WERs, they **must** be fed four hours before each renewal. The test solutions **must** be renewed every 24 hours.

NEC. A 20-day IC25 from a survival and growth renewal test (renew every 48 hours) with a species in the genus Neanthes (Johns et al. 1991). [Note: When determining WERs, sediment **must not** be in the test chamber.] [Note: This test has not been widely used.]

COMMENTS

- X. Another sensitive test cannot be identified at this time, and so other tests used in the criteria document should be considered.
- Y. Because neither the CCCs for mercury nor the freshwater criterion for selenium is based on laboratory data concerning toxicity to aquatic life, they cannot be adjusted using a WER.

REFERENCES

ASTM. 1993a. Guide for Conducting Acute Toxicity Tests with Fishes, Macroinvertebrates, and Amphibians. Standard E729. American Society for Testing and Materials, Philadelphia, PA.

ASTM. 1993b. Guide for Conducting Static Acute Toxicity Tests Starting with Embryos of Four Species of Saltwater Bivalve Molluscs. Standard E724. American Society for Testing and Materials, Philadelphia, PA.

Burgess, R., G. Morrison, and S. Rego. 1992. Standard Operating Procedure for 7-day Static Sublethal Toxicity Tests for Mulinia lateralis. U.S. EPA, Environmental Research Laboratory, Narragansett, RI.

- Chapman, G.A. 1992. Sea Urchin (Strongylocentrotus purpuratus) Fertilization Test Method. U.S. EPA, Newport, OR.
- Johns, D.M., R.A. Pastorok, and T.C. Ginn. 1991. A Sublethal Sediment Toxicity Test using Juvenile Neanthes sp. (Polychaeta:Nereidae). In: Aquatic Toxicology and Risk Assessment: Fourteenth Volume. ASTM STP 1124. (M.A. Mayes and M.G. Barron, eds.) American Society for Testing and Materials, Philadelphia, PA. pp. 280-293.
- Peltier, W.H. 1993. Memorandum to C.E. Stephan. October 19.
- Spehar, R.L., and A.R. Carlson. 1984. Derivation of Site-Specific Water Quality Criteria for Cadmium and the St. Louis River Basin, Duluth, Minnesota. Environ. Toxicol. Chem. 3:651-665.
- U.S. EPA. 1993a. Methods for Measuring the Acute Toxicity of Effluents and Receiving Waters to Freshwater and Marine Organisms. Fourth Edition. EPA/600/4-90/027F. National Technical Information Service, Springfield, VA.
- U.S. EPA. 1993b. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Freshwater Organisms. Third Edition. EPA/600/4-91/002. National Technical Information Service, Springfield, VA.
- U.S. EPA. 1993c. Short-term Methods for Estimating the Chronic Toxicity of Effluents and Receiving Waters to Marine and Estuarine Organisms. Second Edition. EPA/600/4-91/003. National Technical Information Service, Springfield, VA.
- Voyer, R.A., and D.G. McGovern. 1991. Influence of Constant and Fluctuating Salinity on Responses of Mysidopsis bahia Exposed to Cadmium in a Life-Cycle Test. Aquatic Toxicol. 19:215-230.

Appendix J: Recommended Salts of Metals

The following salts are recommended for use when determining a WER for the metal listed. If available, a salt that meets American Chemical Society (ACS) specifications for reagent-grade should be used.

Aluminum

- *Aluminum chloride 6-hydrate: $\text{AlCl}_3 \cdot 6\text{H}_2\text{O}$
- Aluminum sulfate 18-hydrate: $\text{Al}_2(\text{SO}_4)_3 \cdot 18\text{H}_2\text{O}$
- Aluminum potassium sulfate 12-hydrate: $\text{AlK}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$

Arsenic(III)

- *Sodium arsenite: NaAsO_2

Arsenic(V)

- Sodium arsenate 7-hydrate, dibasic: $\text{Na}_2\text{HASO}_4 \cdot 7\text{H}_2\text{O}$

Cadmium

- Cadmium chloride 2.5-hydrate: $\text{CdCl}_2 \cdot 2.5\text{H}_2\text{O}$
- Cadmium sulfate hydrate: $3\text{CdSO}_4 \cdot 8\text{H}_2\text{O}$

Chromium(III)

- *Chromic chloride 6-hydrate (Chromium chloride): $\text{CrCl}_3 \cdot 6\text{H}_2\text{O}$
- *Chromic nitrate 9-hydrate (Chromium nitrate): $\text{Cr}(\text{NO}_3)_3 \cdot 9\text{H}_2\text{O}$
- Chromium potassium sulfate 12-hydrate: $\text{CrK}(\text{SO}_4)_2 \cdot 12\text{H}_2\text{O}$

Chromium(VI)

- Potassium chromate: K_2CrO_4
- Potassium dichromate: $\text{K}_2\text{Cr}_2\text{O}_7$
- *Sodium chromate 4-hydrate: $\text{Na}_2\text{CrO}_4 \cdot 4\text{H}_2\text{O}$
- Sodium dichromate 2-hydrate: $\text{Na}_2\text{Cr}_2\text{O}_7 \cdot 2\text{H}_2\text{O}$

Copper

- *Cupric chloride 2-hydrate (Copper chloride): $\text{CuCl}_2 \cdot 2\text{H}_2\text{O}$
- Cupric nitrate 2.5-hydrate (Copper nitrate): $\text{Cu}(\text{NO}_3)_2 \cdot 2.5\text{H}_2\text{O}$
- Cupric sulfate 5-hydrate (Copper sulfate): $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$

Lead

- *Lead chloride: PbCl_2
- Lead nitrate: $\text{Pb}(\text{NO}_3)_2$

Mercury

- Mercuric chloride: HgCl_2
- Mercuric nitrate monohydrate: $\text{Hg}(\text{NO}_3)_2 \cdot \text{H}_2\text{O}$
- Mercuric sulfate: HgSO_4

Nickel

- *Nickelous chloride 6-hydrate (Nickel chloride): $\text{NiCl}_2 \cdot 6\text{H}_2\text{O}$
- *Nickelous nitrate 6-hydrate (Nickel nitrate): $\text{Ni}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$
- Nickelous sulfate 6-hydrate (Nickel sulfate): $\text{NiSO}_4 \cdot 6\text{H}_2\text{O}$.

Selenium(IV)

- *Sodium selenite 5-hydrate: $\text{Na}_2\text{SeO}_3 \cdot 5\text{H}_2\text{O}$

Selenium(VI)

- *Sodium selenate 10-hydrate: $\text{Na}_2\text{SeO}_4 \cdot 10\text{H}_2\text{O}$

Silver

Silver nitrate: AgNO_3

(Even if acidified, standards and samples containing silver must be in amber containers.)

Zinc

Zinc chloride: ZnCl_2

*Zinc nitrate 6-hydrate: $\text{Zn}(\text{NO}_3)_2 \cdot 6\text{H}_2\text{O}$

Zinc sulfate 7-hydrate: $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$

*Note: ACS reagent-grade specifications might not be available for this salt.

No salt should be used until information concerning the safety and handling of that salt has been read.